

RESULT	15	
014898		
ID	014898	PRELIMINARY; PRT; 506 AA.
AC	014898	
DT	01-JAN-1998	(TREMBLrel. 05, Created)
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)
DT	01-MAY-1999	(TREMBLrel. 10, Last annotation update)
DE	DE	ZINC FINGER PROTEIN (FRAGMENT).
GN	HZF6.	
OS	Homo sapiens (Human).	
JC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.	
KN	[1]	
TX	SEQUENCE FROM N.A.	
TX	MEDLINE; 98234550.	
RA	SHANNON M., STUBBS L.;	
RT	"Analysis of homologous XRCC1-linked zinc-finger gene families in	
RT	human and mouse: evidence for orthologous genes.";	
RL	Genomics 49:112-121(1998).	
DR	EMBL; AF027513; AAD12728.1; -.	
DR	PFAM; PF01352; KRAB; 1.	
DR	PFAM; PF00096; zf-C2H2; 8.	
DR	PROSITE; PS00028; ZINC_FINGER_C2H2; 6.	
KW	zinc-finger; Metal-binding; DNA-binding.	
FT	NON_TER 506	
SQ	SEQUENCE 506 AA; 57498 MW; 7A530D9B CRC32;	


```
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE ZINC-FINGER PROTEIN ZFP-37.
GN ZFP-37.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA/CA, CBA/C57BL, C57BL/6; TISSUE-BRAIN;
RC MEDLINE: 96301401.
PA MAZARAKIS N., MICHALOVICH D., KARIS A., GROSVELD F., GALTART N.;
PT "Zfp-37 is a member of the KRAB zinc finger gene family and is
RT expressed in neurons of the developing and adult CNS.";
RL Genomics 33:247-257(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA/CA, CBA/C57BL, C57BL/6; TISSUE-BRAIN;
RC MEDLINE: 90301500.
PA NELKI D., DUDLEY K., CUNNINGHAM P., AKHAVAN M.;
PT "Cloning and sequencing of a zinc finger cDNA expressed in mouse
RT testis.";
RL Nucleic Acids Res. 18:3655-3655(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA/CA, CBA/C57BL, C57BL/6; TISSUE-BRAIN;
RC MEDLINE: 92310982.
PA BURKE P.S., WOLGEMUTH D.J.;
RT "Zfp-37, a new murine zinc finger encoding gene, is expressed in a
RT developmentally regulated pattern in the male germ line.";
RL Nucleic Acids Res. 20:2827-2834(1992).
DR EMBL: X89264; CAA61539.1; -.
DR PFAM: PF01352; KRAB; 1.
DR PROSITE: PS00096; zf-C2H2; 12.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 11.
KW zinc-finger; Metal-binding; DNA-binding.
FT CONFLICT 101 101 S -> T (IN REF. 3).
FT CONFLICT 300 300 L -> V (IN REF. 3).
FT CONFLICT 528 528 T -> P (IN REF. 3).
FT CONFLICT 533 533 K -> N (IN REF. 3).
FT CONFLICT 572 572 F -> V (IN REF. 2).
SQ SEQUENCE 594 AA; 67254 MW; 42292DA3 CRC32;

Query Match 26.2%; Score 295; DB 11; Length 594;
Best Local Similarity 38.0%; Pred. No. 3.3e-21;
Matches 57; Conservative 29; Mismatches 44; Indels 20; Gaps 3;

QY 48 PTSGK-MNCDVCLSCISFNVLVHRSHTGTPQCNCQCGASFTQKGNLLRHILHTGE 106
DB 249 PUGRPYECNHCCKVLSHKGGLDHTHTGKPYECNCGAFSGKSHLYLHQRTHTGE 308
QY 107 KPFKCHLCNYACQRRDALGHLRTHSVEKPYKCFGRSKYKORSSLEEH----- 156
DB 309 KPYECQCKGAKGHKHALDHLRIHTGEPKPYKCEGKTFRRSSNLMQHLRSHTSKPYE 368
QY 156 -KERCTFLQSDPGDTASAEARHIKAEMG 184
DB 369 CKECGKSFYN-----SSLTEHVTHTG 391

RESULT 9
ID Q15776 PRELIMINARY; PRT; 578 AA.
AC Q15776;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE ZINC FINGER PROTEIN.
GN LD5-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

Query Match 25.8%; Score 291; DB 11; Length 591;
Best Local Similarity 37.3%; Pred. No. 8.1e-21;
Matches 62; Conservative 28; Mismatches 52; Indels 24; Gaps 5;

QY 36 LERHVVSDSRPTSGKMNCDVCLSCISFNVLVHRSHTGTPQCNCQCGASFTQKGN 95
RN [1]
RP SEQUENCE FROM N.A.
RC BELLEFROID E.J., BOURGUIGNON C., SAHIN M., PIELER T., WARD D.C.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U67082; AAC61661.1; -.
DR PFAM: PF01352; KRAB; 1.
DR PROSITE: PS00096; zf-C2H2; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 6.
KW zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 591 AA; 68328 MW; 047D7E62 CRC32;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-OVARY;
RA BEUTLER E., GELBART T., WEST C., KUHL W., LEE P.L.;
RL Blood Cells Mol. Dis. 21:206-216(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-OVARY;
RA LEE P.L.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97386587.
RA LEE P.L., GELBART T., WEST C., ADAMS M., BLACKSTONE R., BEUTLER E.;
RT "Three genes encoding zinc finger proteins on human chromosome
RT 6p21.3: members of a new subclass of the Kruppel gene family
RT containing the conserved SCAN box domain.";
RL Genomics 43:191-201(1997).
DR EMBL: U57796; AAB02260.1; -.
DR EMBL: U88080; AAC51656.1; -.
DR PFAM: PF00096; zf-C2H2; 9.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 9.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 578 AA; 65757 MW; C0A9AD2B CRC32;

Query Match 25.9%; Score 291.5; DB 4; Length 578;
Best Local Similarity 40.3%; Pred. No. 7e-21;
Matches 56; Conservative 24; Mismatches 54; Indels 5; Gaps 2;

QY 19 EIPYSYREYNEY-ENIKLERHVVFDSPRTSGKMNCDVCLSCISFNVLVHRSHTG 77
DB 431 ERPY-----ECNECGKAFSSHLLGHQRIHTGEPKPYECDECGKTFRRSSHLIGHORSHTG 486
QY 78 ERPQCNCQCGASFTQKGNLLRHILHTGEPKFKCHLCNYACQRRDALGHLRTHSVEKPY 137
DB 487 EKPYKCEGKRAFSKGLIEHQRIHTGEPKPYKCEGKAFNGNTGLIQLHRIHTGKPY 546
QY 138 KCEFCGRSKYKORSSLEEHK 156
DB 547 QCNECGRAFIQSSLIHQ 565

RESULT 10
ID P70590 PRELIMINARY; PRT; 591 AA.
AC P70590;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE KRAB-ZINC FINGER PROTEIN KZF-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA BELLEFROID E.J., BOURGUIGNON C., SAHIN M., PIELER T., WARD D.C.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U67082; AAC61661.1; -.
DR PFAM: PF01352; KRAB; 1.
DR PROSITE: PS00096; zf-C2H2; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 6.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 591 AA; 68328 MW; 047D7E62 CRC32;

Query Match 25.8%; Score 291; DB 11; Length 591;
Best Local Similarity 37.3%; Pred. No. 8.1e-21;
Matches 62; Conservative 28; Mismatches 52; Indels 24; Gaps 5;

QY 36 LERHVVSDSRPTSGKMNCDVCLSCISFNVLVHRSHTGTPQCNCQCGASFTQKGN 95
RN [1]
RP SEQUENCE FROM N.A.
RC BELLEFROID E.J., BOURGUIGNON C., SAHIN M., PIELER T., WARD D.C.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U67082; AAC61661.1; -.
DR PFAM: PF01352; KRAB; 1.
DR PROSITE: PS00096; zf-C2H2; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 6.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 591 AA; 68328 MW; 047D7E62 CRC32;
```

```

OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98056818.
RA HANSEN J.D., STRASSBURGER P., DU PASQUIER L.;
RT "Conservation of a master hematopoietic switch gene during vertebrate
RT evolution: isolation and characterization of ikaros from teleost and
RT amphibian species.";
KL Eur. J. Immunol. 27:3049-3058(1997).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER
CC (DELTA-A ELEMENT) OF THE CD3-DELTA GENE.
CC FUNCTIONS IN THE SPECIFICATION AND THE MATURATION OF THE T
CC LYMPHOCYTE.
CC ALSO INTERACTS WITH A CRITICAL CONTROL ELEMENT IN THE TDT
CC (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE) PROMOTER AS WELL AS WITH
CC THE PROMOTERS FOR OTHER GENES EXPRESSED DURING EARLY STAGES OF B
CC AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THYMUS AND SPLEEN.
CC NO EXPRESSION IN MUSCLE.
CC -!- SIMILARITY: TO D.MELANOGASTER HUNCHBACK.
DR EMBL; U92202; AAB53435.1; -.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
DR PFAM; PF00096; zf-C2H2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein.
FT NON_TER 1 1
FT ZN_FING 5 27 C2H2-TYPE.
FT ZN_FING 33 55 C2H2-TYPE.
FT ZN_FING 61 84 C2H2-TYPE.
FT NON_TER 138 138
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15701 MW; 2B786CAD CRC32;

Query Match 48.6%; Score 547.5; DB 13; Length 138;
Best Local Similarity 75.4%; Pred. No. 9.8e-47;
Matches 104; Conservative 9; Mismatches 18; Indels 7; Gaps 1;

QY 77 GERPFQCCGASFTQKGNLLRHILKHTGERPFKCHLCNYACORRDALTGHLRTHSVKPK 136
DB 1 GERPFQCCGASFTQKGNLLRHILKHTGERPFKCHLCNYACORRDALTGHLRTHSVKPK 60

QY 137 YKCEFCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM-----GSEAL 189
DB 137 YKCEFCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM-----GSEAL 189

QY 61 HKCGYCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM-----GSEAL 120
DB 61 HKCGYCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM-----GSEAL 120

QY 190 VLDRLASNVAKRKSSMPQ 207
DB 121 VLDRLASNVAKRKSSMPQ 138

Query Match 26.5%; Score 298.5; DB 11; Length 346;
Best Local Similarity 35.1%; Pred. No. 8.3e-22;
Matches 59; Conservative 34; Mismatches 68; Indels 7; Gaps 3;

ID 042244 PRELIMINARY; PRT; 328 AA.
AC 042244;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE IKAROS-RELATED TRANSCRIPTION FACTOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RA TURPEN J., KELLEY C., MEAD P., ZON L.;
RL Immunity 0:0-0(1997).
DR EMBL; AF024439; AAB81280.1; -.
DR PFAM; PF00096; zf-C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 2.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1 1
FT NON_TER 328 328

```

```

SQ SEQUENCE 328 AA; 37116 MW; CCC9E71E CRC32;

Query Match 30.9%; Score 347.5; DB 13; Length 328;
Best Local Similarity 55.3%; Pred. No. 1.2e-26;
Matches 68; Conservative 13; Mismatches 27; Indels 15; Gaps 2;

QY 102 LHTGEKPFKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEHKRCRT 161
DB 102 LHTGEKPFKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEHKRCRT 161

QY 162 FLQS-----TDPGDTASAEARHIKAEMG---SERALVLDRLASNVAKRKSSMP 206
DB 162 FLQS-----TDPGDTASAEARHIKAEMG---SERALVLDRLASNVAKRKSSMP 206

QY 61 YLQNVMEAGQIVHHASPMDECKNPETMMNMSLMPFERPAVIERLASNMGRKKSSTP 120
DB 61 YLQNVMEAGQIVHHASPMDECKNPETMMNMSLMPFERPAVIERLASNMGRKKSSTP 120

QY 207 QKF 209
DB 121 QRF 123

RESULT 7
Q62512 ID Q62512 PRELIMINARY; PRT; 346 AA.
AC Q62512;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE ZINC FINGER PROTEIN (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA FUJIWARA Y.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA FUJIWARA Y.;
RL "Studies on zinc-finger protein genes expressed in mouse
RL spermatogenesis.";
RL Thesis (1991), Unknown Institution.
DR EMBL; D10627; BAA01477.1; -.
DR PFAM; PF00096; zf-C2H2; 12.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 12.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1 1
FT NON_TER 346 346
SQ SEQUENCE 346 AA; 39665 MW; 06C78AAD CRC32;

Query Match 26.5%; Score 298.5; DB 11; Length 346;
Best Local Similarity 35.1%; Pred. No. 8.3e-22;
Matches 59; Conservative 34; Mismatches 68; Indels 7; Gaps 3;

QY 16 EPEIPYSYGREYNEYENIKLERHVVSDSSRPTSGKMNCDVGLSCISFNVLVHKRSH 75
DB 16 EPEIPYSYGREYNEYENIKLERHVVSDSSRPTSGKMNCDVGLSCISFNVLVHKRSH 75

QY 76 TGERPFQCCGASFTQKGNLLRHILKHTGERPFKCHLCNYACORRDALTGHLRTHSVK 135
DB 76 TGERPFQCCGASFTQKGNLLRHILKHTGERPFKCHLCNYACORRDALTGHLRTHSVK 135

QY 114 TGEYPYECNECGKAFSQHSSQLQNHKRTHTGKPYECNECGKAFSQHSSQLQNHKRTHTG 173
DB 114 TGEYPYECNECGKAFSQHSSQLQNHKRTHTGKPYECNECGKAFSQHSSQLQNHKRTHTG 173

QY 136 PYKCEFCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM 183
DB 136 PYKCEFCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM 183

QY 174 PYECNECGKAFSQHSSQLQNHKRTHTGKPYECNECGKAFSQHSSQLQNHKRTHTG 220
DB 174 PYECNECGKAFSQHSSQLQNHKRTHTGKPYECNECGKAFSQHSSQLQNHKRTHTG 220

RESULT 8
Q62514 ID Q62514 PRELIMINARY; PRT; 594 AA.
AC Q62514;
DT 01-NOV-1996 (Tremblrel. 01, Created)

```


DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE IKAROS.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 OC Cyprinoidae; Cyprinidae; Rasbora; Danio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA AMEMIYA C., KAWASAKI H.;
 RT "Characterization of zebrafish ikaros, a gene necessary for
 RT differentiation of the immune system.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF092175; AAC61763.1;
 DR PFAM: PF00096; zf-C2H2; 5.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 KW zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 537 AA; 58865 MW; 6AF8330F CRC32;

 Query Match 57.7%; Score 650; DB 13; Length 537;
 Best Local Similarity 55.7%; Pred. No. 3.5e-56;
 Matches 137; Conservative 17; Mismatches 38; Indels 54; Gaps 6;

 QY 3 DENVLSEPMGNAREPEI-----PYSREYNEVENIKLERHVVSF 43
 DB 66 EENGLSCENGEAECAEDRLDLSGAKVNGSHAGPDSKPAAYTAGIRL-----119

 QY 44 DSSRPTSGKNCDCVGLSCISFNVLVHKRSHT-----GERPQCNOCGASFTQ 92
 DB 119 -----PGLKLCIDICIGVIGPNVLVHKRSHTPEERKSVLEQKQGERPQCNOCGASFTQ 173

 QY 93 KGNLRHKLHTGEKPKFCHLCNYACQRRDALTGLHRTSHVEPKYKCEFCGRSYKORSSL 152
 DB 174 KGNLRHKLHSGEKFCHLCNYACRRDALTGLHRTSHVGPKCAICGRSYKORSSL 233

 QY 153 EEKERCRTFLOS-----TPDGTASAEARHIAEM-GSERALVLDRLASNAVKRKS 203
 DB 234 EEKERCRTFLOS-----TPDGTASAEARHIAEM-GSERALVLDRLASNAVKRKS 290

 QY 204 SMPQKF 209
 DB 291 SMPQKF 296

 RESULT 3
 C92222 PRELIMINARY; PRT; 533 AA.
 AC Q92222
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE PPS PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR;
 RA HOMMA Y., KIYOSAWA H., MORI T., OGURI A., NIKAIKO T., KANAZAWA K.,
 RA TOJO M., TAKEDA J., TANNI Y., YOKOYA S., KAWABATA I., IKEDA H.,
 RA WANAKA A.;
 RT "Eos: a novel member of the Ikaros gene family expressed
 RT predominantly in the developing nervous system.";
 RL FEBS Lett. 0:0-0(1999).
 DR EMBL: AB017615; BAA36213.1;
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 533 AA; 58167 MW; 2A0315AD CRC32;

 Query Match 54.8%; Score 617; DB 11; Length 533;

Best Local Similarity 53.6%; Pred. No. 6.1e-53;
 Matches 127; Conservative 23; Mismatches 33; Indels 54; Gaps 8;

 QY 1 ERDENVL---KSEPM---GNAEPEIPYYSREYNEVENIKLERHVVSPDSSRPTSGK 52
 DB 68 DKDDSVIVEDSLSEPLGCDGSGPEHPSP-----GGIRLPGNK 105

 QY 53 MNCDCVGLSCISFNVLVHKRSHTGERPQCNOCGASFTQKGNLRHKLHTGERKPKCH 112
 DB 106 LKCDVCGMVICGPNVLVHKRSHTGERPQCNOCGASFTQKGNLRHKLHSGERPKCP 165

 QY 113 LONACQRRDALTGLHRTSHVE-----KPYKCEFCGRSYKORSSLEEKERCRTFLO--S 165
 DB 166 FCNYACRRDALTGLHRTSHVSSPTVGPKYCNCGRSYKQOSTLEEKERCRTFLOSL 225

 QY 166 TD-----PGDTASAEARHIAEMGSEALVLDRLASNAVKRKSMPQKF 209
 DB 226 TDAQALTQPGD-----EIDLEWVPSMLHPSTERTFIDRLANSLTKRKRSTPQKF 278

 RESULT 4
 P79751 PRELIMINARY; PRT; 417 AA.
 AC P79751
 DT 01-MAY-1997 (TReMBLrel. 03, Created)
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE IKAROS-LIKE.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
 OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA GELLNER K., BRENNER S.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF056116; AAC34387.1;
 DR PFAM: PF00096; zf-C2H2; 5.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 417 AA; 45664 MW; 7856E12F CRC32;

 Query Match 51.4%; Score 578.5; DB 13; Length 417;
 Best Local Similarity 62.0%; Pred. No. 2.9e-49;
 Matches 103; Conservative 25; Mismatches 35; Indels 3; Gaps 1;

 QY 47 RPTSGKNCDCVGLSCISFNVLVHKRSHTGERPQCNOCGASFTQKGNLRHKLHTGE 106
 DB 112 RLPNGKLOCEICGIVCTGPNVLVHKRSHTGERPQCNOCGASFTQKGNLRHKLHSGE 171

 QY 107 KPFKCHLCNYACQRRDALTGLHRTSHVEPKYKCEFCGRSYKORSSLEEKERCRTFLOST 166
 DB 172 KPFKCPICNYACRRDALTGLHRTSHVGPKCAICGRSYKQOSTLEEDHQRCHSYLKR 231

 QY 167 DPGDTASAEARHIAEMGSEALVLDRLASNAVKRKSMPQKF 209
 DB 232 QAAVVRQAAPGALNMDGMNOSNEKIQOVDRLANTIAKRRTMPQKF 277

 RESULT 5
 O13100 PRELIMINARY; PRT; 138 AA.
 AC O13100
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1)
 DE (FRAGMENT).
 GN IKAROS OR LYF-1.
 OS Xenopus laevis/gilli.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;

Result No.	Score	Query Match	Length	ID	Description
1	1024.5	91.0	507	11	O08900 mus musculus
2	650	57.7	537	13	O93581 brachydanio
3	617	54.8	533	11	O92222 mus musculus
4	578.5	51.4	417	13	P79751 fugu rubripes
5	547.5	48.6	138	13	Q13100 xenopus laevis
6	547.5	30.9	328	13	Q42244 xenopus laevis
7	298.5	26.5	346	11	O62512 mus musculus
8	295	26.2	594	11	O62514 mus musculus
9	291.5	25.9	578	4	Q15776 homo sapien
10	291	25.8	591	11	P70590 rattus norv
11	289	25.7	819	11	Q921d8 mus musculus
12	289	25.7	292	11	Q92005 mus musculus
13	288	25.6	814	11	O70162 mus musculus
14	287	25.5	367	4	Q13398 homo sapien
15	286.5	25.4	506	4	O14898 homo sapien
16	286	25.4	605	4	O60765 homo sapien
17	286	25.4	650	11	O62886 rattus norv
18	285	25.3	1167	4	O43345 homo sapien
19	284.5	25.3	612	4	O75802 homo sapien
20	284	25.2	732	4	Q14586 homo sapien
21	284	25.2	395	4	O43693 homo sapien
22	284	25.2	430	4	O95015 homo sapien
23	284	25.2	812	11	Q35483 mus musculus
24	283.5	25.2	1207	4	O43724 homo sapien
25	281.5	25.0	427	4	Q02313 homo sapien
26	280	24.9	385	4	O00455 homo sapien
27	279.5	24.8	751	4	O60792 homo sapien
28	279	24.8	445	4	O95779 homo sapien
29	279	24.8	498	4	O95780 homo sapien

DT 01-NOV-1998 (TREMBlrel. 08; Last sequenced)

DT 0

08; Last sequence up

DT 01-NOV-1998 (TREMBLrel: 08; Last sequence update)

Search completed: November 6, 1999, 11:46:21
Job time: 6525 sec

```

FT  2N_FING      268      290      C2H2-TYPE.
FT  2N_FING      296      318      C2H2-TYPE.
FT  2N_FING      324      346      C2H2-TYPE.
SQ  SEQUENCE     347 AA: 39855 MW;  F042FE59  CRC32;

      Query Match      25.8%      Score 290:  DB 1:  Length 347;
      Best Local Similarity 34.8%;  Pred. No. 5.2e-19;
      Matches 54;  Conservative 34;  Mismatches 61;  Indels 6;  Gaps 2;

QY  2  RDNVLKSEPMGNAEEPIPYSYSREYENIENIKLERHVVFSRPTSGRMNCDVCGLS 61
    :  :||| :  :||| :  :||| :  :||| :  :||| :
Db  111 QSSLLKHQRHITGKPYTCVCDKHFIERSLTV--HQRHTGKPK---YKCECGKA 164

QY  62  CISENVLMVKRSHTGRRPFQCNCGASFTQGNLLRHILKLTGKEPKFKCHLQNYACORR 121
    :  :||| :  :||| :  :||| :  :||| :  :||| :
Db  165 FSQSMNLTVHQRHTGKPYCKCKGKAFRNSLSIQHERHTGKPYKCHDCGKAFSKN 224

QY  122 DALTGHLRTHSVEKPYKCFECGRSYKORSSLEHK 156
    :  :||| :  :||| :  :||| :
Db  225 SSLTQHRRIITGKPYECMIGCKHFTGRSSLTVHQ 259
    :  :||| :  :||| :  :||| :

```

```

RESULT 13
OZF_BOVIN
ID OZF_BOVIN STANDARD; PRT; 292 AA.
AC Q28151;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN OZF.
GN ZNF146 OR OZF.
OS BOA TAUROS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.
HN [1]
RP SEQUENCE FROM N.A.
AC
AC ISSUE=LUNG;
AC MEDLINE; 96152243.
RA LE CHALONY C., APIOU F., PIBOUIN L., DUTRILLAUX B., GOUBIN G.;
RT "Constitutive amplification of a zinc finger protein gene in cattle.";
RL DNA CELL BIOL. 15:83-88(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE KRUEPEL SUBFAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/)
or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; X81804; E213761; ALT_INIT.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 10.
DR PFAM; PF00096; zf-C2H2; 10.
DR HSP; P25490; 12NM.
KW ZINC-FINGER; DNA-BINDING; METAL-BINDING; NUCLEAR PROTEIN; REPEAT.
FT DOMAIN 16 292 ZINC-FINGERS.
FT ZN_FING 16 38 C2H2-TYPE.
FT ZN_FING 44 66 C2H2-TYPE.
FT ZN_FING 72 94 C2H2-TYPE.
FT ZN_FING 100 122 C2H2-TYPE.
FT ZN_FING 128 150 C2H2-TYPE.
FT ZN_FING 156 178 C2H2-TYPE.
FT ZN_FING 184 206 C2H2-TYPE.
FT ZN_FING 212 234 C2H2-TYPE.
FT ZN_FING 240 262 C2H2-TYPE.
FT ZN_FING 268 292 C2H2-TYPE.
SQ SEQUENCE 292 AA; 33269 MW; 88B8A3DD CRC32.

```

```

Query Match          25.7%; Score 289; DB 1; Length 292;
Best Local Similarity 30.4%; Pred. No. 5.3e-19;
Matches 62; Conservative 45; Mismatches 67; Indels 30; Gaps 5;

Qy 4 ENVLKSPGMNAEPEIPYPSYSREYNEYENIKLHRHVVSFDSSRP-----TSGKMN 54
      |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 85 ENLLTHOKIHGKPFCKDCGKAFIQKSN--LIRHQRTHTGKPFCKGCKTFSQKSN 142
      |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 55 -----CDVCGLSCLISFNVLVHKRSHTGERPFCQNCOCGASFTQKGNLLRH 99
      |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 143 LTEHEKTHIGKPKFQNECGTAFQCKYLYLKQNIHGTGKPYECNECGKAFSQTSLIVH 202
      |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 100 IKLTGKPKFKCHLCNYACQRRDALTLGHLRTHSVKPYKCFGCGRSYKQKRSLSLEHKERC 159
      ::||: |||: |||: : : : : : : : : : : : : : : : : :
Db 203 VRIHSGDKPYECNCGKAFSGSSSLTVHRSHTGKPYGCGNECGKAFSQFSTLALHL-RI 261
      |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 160 RTF---LQSTDPDGDTASAEARHIK 180
      | : : : : : : : : : : : : : : : : : : : : : : :
Db 262 HTGRKPYQCECGKAFSQKSHHIR 285
      | : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Z141_HUMAN STANDARD; PRT; 474 AA.
ID Z141_HUMAN AC Q15928;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 141.
GN ZNF141.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
RN PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RC [1]
RP SEQUENCE FROM N.A.
RX TISSUE-INSULINOMA;
RC MEDLINE; 94093541.
RA TOMMERUP N., AAGAARD L., LUND C.L., BOEL E., BAXENDALE S., BATES G.P.,
RA LEHRACH H., VISSING H.;
RT "A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate
RT gene for the Wolf-Hirschhorn (4p-) syndrome.";
RL HUM. MOL. GENET. 2:1571-1575(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION AS A
CC REPRESSOR. CANDIDATE GENE FOR WOLF-HIRSCHHORN (4p-) SYNDROME
CC (WHS).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL);
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY LOW EXPRESSION.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L15309; G347906; -.
CC MIM; 194648; -.
DR DR
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 10.
DR PFAM; PF00096; zf-C2H2; 10.
DR HSPF; P08047; 15P2.
KW TRANSCRIPTION REGULATION; REPRESSOR; ZINC-FINGER; METAL-BINDING;
KW NUCLEAR PROTEIN.
FT DOMAIN 2 43 KRAB BOX ("A BOX").
FT DOMAIN 44 75 KRAB BOX ("B BOX").
FT DOMAIN 199 473 ZINC-FINGERS.
FT ZN_FING 199 221 C2H2-TYPE.
FT ZN_FING 227 249 C2H2-TYPE.
FT ZN_FING 255 277 C2H2-TYPE.
FT ZN_FING 283 305 C2H2-TYPE.
FT FT

```



```

ZN_FING 107 129 C2H2-TYPE.
ZN_FING 135 157 C2H2-TYPE.
ZN_FING 163 185 C2H2-TYPE.
ZN_FING 191 213 C2H2-TYPE.
ZN_FING 236 258 C2H2-TYPE.
ZN_FING 264 286 C2H2-TYPE.
ZN_FING 292 314 C2H2-TYPE.
ZN_FING 320 342 C2H2-TYPE.
ZN_FING 348 370 C2H2-TYPE.
ZN_FING 376 398 C2H2-TYPE.
ZN_FING 404 426 C2H2-TYPE.
ZN_FING 432 454 C2H2-TYPE.
ZN_FING 460 482 C2H2-TYPE.
SEQUENCE 485 AA; 54463 MW; 908B33F1 CRC32;

Query Match 26.4%; Score 297; DB 1; Length 485;
Best Local Similarity 41.0%; Pred. No. 1.8e-19;
Matches 57; Conservative 22; Mismatches 50; Indels 10; Gaps 2;

ZY QY 18 PEIPYSRYEYENIKLERHVVSFDSSRPTSGKMKNCVCGLSISFNVLWVKRSHTG 77
DDB 351 PECQRF-----QRKLTRHORHTHGKP----YHCGEGLGTQVSRTEHQRIHTG 400
ZY QY 78 ERPFCQCQGASFTOKGNLLRHILKTGKPKFKCHLYACQRDALTGHLRTHSVEKPY 137
DDB 401 ERPEACPCEGOSPROHANLTOHRRIHTGERPYACPEGKAFQRPTLTQHRLTRRREKPF 460
ZY QY 138 KCFGRSYKQRSLEHK 156
DDB 451 ACDCGRGFHQSTKLIOHQ 479

RESULT 9
ZN91_HUMAN STANDARD; PRT; 1191 AA.
IID ZN91_HUMAN Q05481;
AC Q05481;
01-JUN-1994 (REL. 29, CREATED)
01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
ZINC FINGER PROTEIN 91 (ZINC FINGER PROTEIN HTF10) (HPF7).
ZN91.
HOMO SAPIENS (HUMAN).
EUARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
RN SEQUENCE FROM N.A.
MEDLINE; 93223677.
RA BELLEFROID E.J., MARINE J.C., RIED T., LECOCQ P.J., RIVIERE M.,
RA AMEMIRA C.T., PONCELET D.A., COULIE P.G., DE JONG P.J.,
RA SPIRER C., WARD D.C., MARTIAL J.A.;
"Clustered organization of homologous KRAB zinc-finger genes with
enhanced expression in human T lymphoid cells.";
RNL EMBO J. 12:1363-1374(1993).
[RN [2]
SEQUENCE OF 15-204 FROM N.A.
MEDLINE; 91219421.
RX BELLEFROID E.J., PONCELET D.A., LECOCQ P.J., REVELANT O.,
RX MARTIAL J.A.;
"the evolutionarily conserved Kruppel-associated box domain defines a
subclass of eukaryotic multifingered proteins.";
RNL PROC. NATL. ACAD. SCI. U.S.A. 88:3608-3612(1991).
CCC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBALE).
CCC -1- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KRAB BOX.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
DR PDB: 1ZNF: 15-OCT-91.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 35.
DR PFAM: PF00096; zf-C2H2; 36.
KW ZINC-FINGER; METAL-BINDING; RNA-BINDING; REPEAT; 3D-STRUCTURE;
KW PHOSPHORYLATION.
FT DOMAIN 1 26 KRAB BOX ("A BOX").
FT DOMAIN 27 58 KRAB BOX ("B BOX").
FT DOMAIN 108 298 ZINC-FINGERS I.
FT DOMAIN 326 488 ZINC-FINGERS II.
FT DOMAIN 503 721 ZINC-FINGERS III.
FT DOMAIN 750 940 ZINC-FINGERS IV.
FT DOMAIN 988 1066 ZINC-FINGERS V.
FT DOMAIN 1136 1298 ZINC-FINGERS VI.
FT ZN_FING 108 130 C2H2-TYPE.
FT ZN_FING 136 158 ZN_FING.
FT ZN_FING 164 186 C2H2-TYPE.
FT ZN_FING 192 214 C2H2-TYPE.
FT ZN_FING 220 242 C2H2-TYPE.
FT ZN_FING 248 270 C2H2-TYPE.
FT ZN_FING 276 298 C2H2-TYPE.
FT ZN_FING 326 348 C2H2-TYPE.
FT ZN_FING 354 376 C2H2-TYPE.
FT ZN_FING 382 404 C2H2-TYPE.
FT ZN_FING 410 432 C2H2-TYPE.
FT ZN_FING 438 460 C2H2-TYPE.
FT ZN_FING 466 488 C2H2-TYPE.
FT ZN_FING 503 525 C2H2-TYPE.
FT ZN_FING 531 553 C2H2-TYPE.
FT ZN_FING 559 587 C2H2-TYPE.
FT ZN_FING 587 609 C2H2-TYPE.
FT ZN_FING 615 637 C2H2-TYPE.
FT ZN_FING 643 665 C2H2-TYPE.
FT ZN_FING 671 693 C2H2-TYPE.
FT ZN_FING 699 721 C2H2-TYPE.
FT ZN_FING 750 772 C2H2-TYPE.
FT ZN_FING 778 800 C2H2-TYPE.
FT ZN_FING 806 828 C2H2-TYPE.
FT ZN_FING 834 856 C2H2-TYPE.
FT ZN_FING 862 884 C2H2-TYPE.
FT ZN_FING 890 912 C2H2-TYPE.
FT ZN_FING 918 940 C2H2-TYPE.
FT ZN_FING 988 1010 C2H2-TYPE.
FT ZN_FING 1016 1038 C2H2-TYPE.
FT ZN_FING 1044 1066 C2H2-TYPE.
FT ZN_FING 1136 1158 C2H2-TYPE.
FT ZN_FING 1164 1186 C2H2-TYPE.
FT ZN_FING 1192 1214 C2H2-TYPE.
FT ZN_FING 1220 1242 C2H2-TYPE.
FT ZN_FING 1248 1270 C2H2-TYPE.
FT ZN_FING 1276 1298 C2H2-TYPE.
FT STRAND 1045 1045
FT STRAND 1052 1052
FT HELIX 1056 1065
FT TURN 1066 1066
SQ SEQUENCE 1350 AA; 155804 MW; 10CA7C8C CRC32;

Query Match 28.2%; Score 317.5; DB 1; Length 1350;
Best Local Similarity 36.9%; Pred. No. 7.6e-21;
Matches 66; Conservative 33; Mismatches 59; Indels 21; Gaps 4;

QY 4 ENVLKSEPMGNAEEPIPYYSREYNEVENIKLRHVVSFDSSRPTSGKMN---CDVGL 60
Db 308 DSVGTDPLSSQVASSPYSCSKRKTFRWK-----SFLNHQOHTSRKPYLCSHCNK 361
QY 61 SCISFNVLVHKRSTGTPFCNCGGASFTQKGNLLRHLKLTGKTFKCHLCLNYACQR 120
Db 362 GFQNSDLVKHRTHTGTPYQACCHGFKIQKSLVKHLRTHTGKTFKCHCDKTFE 421
QY 121 RDALTGHLRTHSEVKPYCEFCGRKYKRSSEEHKE-----RC-----RTFLQSTD 167
Db 422 RSALAKHQHTHGERPKYKSCDCKEFTQRSNLLHQRIHTGTPKCYKCLDRTFIQNSD 480
```

```
RESULT 7
ZN43_HUMAN
AC P28160; STANDARD; PRT; 803 AA.
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 43 (ZINC PROTEIN HTF6).
GN ZNF43.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T-CELL;
RX MEDLINE: 91279444.
RA LOVERING R., TROWSDALE J.;
RT "A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell lines.";
RL NUCLEIC ACIDS RES. 19:2921-2927(1991).
[2]
RP SEQUENCE OF 38-190 FROM N.A.
RX MEDLINE: 91219421.
RA BELLEFROID E.J., PONCELET D.A., LECOCQ P.J., REVELANT O., MARTIAL J.A.;
RT "The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins.";
RL PROC. NATL. ACAD. SCI. U.S.A. 88:3608-3612(1991).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: T AND B CELL LINES.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59244; G38032;
DR EMBL: M61869; G184450;
DR PIR: S26823; S26823.
DR PIR: D39384; D39384.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 19.
DR PFAM: PF00096; zf-C2H2; 21.
DR HSSP: P25490; 1ZNF.
KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN.
FT DOMAIN 1 37 KRAB BOX ("A BOX").
FT DOMAIN 38 69 KRAB BOX ("B BOX").
FT DOMAIN 168 778 22 C2H2-TYPE ZINC-FINGERS.
FT ZN_FING 168 190 C2H2-TYPE.
FT ZN_FING 196 218 C2H2-TYPE (DEGENERATE).
FT ZN_FING 224 246 C2H2-TYPE (DEGENERATE).
FT ZN_FING 252 274 C2H2-TYPE (DEGENERATE).
FT ZN_FING 280 302 C2H2-TYPE.
FT ZN_FING 308 330 C2H2-TYPE.
FT ZN_FING 336 358 C2H2-TYPE.
FT ZN_FING 364 386 C2H2-TYPE.
FT ZN_FING 392 414 C2H2-TYPE.
FT ZN_FING 420 442 C2H2-TYPE.
FT ZN_FING 448 470 C2H2-TYPE.
FT ZN_FING 476 498 C2H2-TYPE.
FT ZN_FING 504 526 C2H2-TYPE.
FT ZN_FING 532 554 C2H2-TYPE.
FT ZN_FING 560 582 C2H2-TYPE.
FT ZN_FING 610 638 C2H2-TYPE.
FT ZN_FING 616 638 C2H2-TYPE.
FT ZN_FING 644 666 C2H2-TYPE.
FT ZN_FING 672 694 C2H2-TYPE.
```



```

FT VARSPLIC 10 53 MISSING (IN IK4).
FT VARSPLIC 54 140 MISSING (IN IK2).
FT VARSPLIC 54 283 MISSING (IN IK6).
FT VARSPLIC 197 283 MISSING (IN IK3 AND IK4).
FT VARSPLIC 141 283 MISSING (IN IK5).
FT CONFLICT 11 12 QV -> FS (IN REF. 2).
FT CONFLICT 214 214 S -> T (IN REF. 2).
FT CONFLICT 245 245 N -> K (IN REF. 2).
FT CONFLICT 296 296 MISSING (IN REF. 2).
FT CONFLICT 298 298 S -> T (IN REF. 2).
FT CONFLICT 352 355 KPLA -> RRS (IN REF. 2).
FT CONFLICT 372 372 N -> Y (IN REF. 2).
FT CONFLICT 420 426 PHARNGL -> RRAQV (IN REF. 2).
SQ SEQUENCE 519 AA; 57528 MW; 5A97272D CRC32;

Query Match 60.8%; Score 684.5; DB 1; Length 519;
Best Local Similarity 74.4%; Pred. No. 4e-54; 20; Indels 9; Gaps 2;
Matches 128; Conservative 15; Mismatches 15;

QY 47 RPTSGKMNCDVGLSCISFNVLVHVKRSHTGERPFCNQCASFTQKGNLLRHKLHTGE 106
DB 111 RLPNGKLKDCIGIICGNPNVLMVHVKRSHTGERPFCNQCASFTQKGNLLRHKLHSGE 170

QY 107 KPFKCHLCNYACORDALTGHLRTHSVKPYKCFGCGRSYKORSLEEHKRCRTFLQST 166
DB 171 KPFKCHLCNYACRRDALTGHLRTHSVGKPHKCGYCGRSYKORSLEEHKRCRCHYLESM 230

QY 167 D-PG-----DTASAEARHKAEMGSRALVLDRLASNVAKRKSSMPQKF 209
DB 231 GLPGTLVPYKTEETHNSEMAEDLCKIGSERSLVLDRLASNVAKRKSSMPQKF 282

RESULT 2
IKAR_CHICK
ID IKAR_CHICK STANDARD; PRT; 518 AA.
AC 042410:
U1 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE DNA-BINDING PROTEIN IKAROS.
GN IKAROS OR IK.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP "Avian Ikaros gene is expressed early in embryogenesis.";
RX STRAIN-H.B2: TISSUE-THYMUS;
RX MEDLINE: 97439462.
RA LIIPPO J., LASSILA O.;
RT EUR. J. IMMUNOL. 27:1853-1857(1997).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES
CC EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC HEMATOPOIETIC ORGANS
CC SUCH AS THE BURSA OF FABRICIUS, THYMUS AND SPLEEN. IN THE ADULT,
CC EXPRESSED IN SPLEEN, THYMUS, BURSA AND PERIPHERAL BLOOD
CC LEUKOCYTES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYO FROM EMBRYONIC DAY
CC 2 ONWARDS.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y11833; E309211; -
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
DR PFAM; PF00096; zf-C2H2; 5.
DR TRANSRIPT; TRANSRIPT; ZINC-FINGER; ZINC-FINGER; METAL-BINDING;
KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
FT DOMAIN 117 224 ZINC-FINGERS I.
FT ZN_FING 461 513 ZINC-FINGERS II.
FT ZN_FING 117 139 C2H2-TYPE.
FT ZN_FING 145 167 C2H2-TYPE.
FT ZN_FING 173 195 C2H2-TYPE.
FT ZN_FING 201 224 C2H2-TYPE.
FT ZN_FING 461 483 C2H2-TYPE.
FT ZN_FING 489 513 C2H2-TYPE.
SQ SEQUENCE 518 AA; 57586 MW; AB4DF123 CRC32;

Query Match 60.1%; Score 676.5; DB 1; Length 518;
Best Local Similarity 73.3%; Pred. No. 2.1e-53;
Matches 126; Conservative 14; Mismatches 23; Indels 9; Gaps 1;

QY 47 RPTSGKMNCDVGLSCISFNVLVHVKRSHTGERPFCNQCASFTQKGNLLRHKLHTGE 106
DB 111 RLPNGKLKDCIGIICGNPNVLMVHVKRSHTGERPFCNQCASFTQKGNLLRHKLHSGE 170

QY 107 KPFKCHLCNYACORDALTGHLRTHSVKPYKCFGCGRSYKORSLEEHKRCRTFLQ-- 165
DB -171 KPFKCHLCNYACRRDALTGHLRTHSVGKPHKCGYCGRSYKORSLEEHKRCRCHYLOTM 230

QY 165 -----STDPGDTASAEARHKAEMGSRALVLDRLASNVAKRKSSMPQKF 209
DB 231 SISSNLYSVKEETNQSEMAEDLCKIGSERSLVLDRLASNVAKRKSSMPQKF 282

RESULT 3
IKAR_ONCMY
ID IKAR_ONCMY STANDARD; PRT; 522 AA.
AC 013089; 013088; 013086; 013087;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS.
GN IKAROS.
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIARDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;
OC SALMONIDAE; ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SHASTA: TISSUE-THYMOCYTES;
RX MEDLINE: 98056818.
RA HANSEN J.D., STRASSBURGER P., DU PASQUIER L.;
RT "Conservation of a master hematopoietic switch gene during vertebrate
RT evolution: isolation and characterization of Ikaros from teleost and
RT amphibian species.";
RL EUR. J. IMMUNOL. 27:3049-3058(1997).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES
CC EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: EIGHT FORMS OF THE PROTEIN, IK1-IK8; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT
CC OF IK1.
CC -!- TISSUE SPECIFICITY: EXPRESSION MAINLY LIMITED TO THYMUS, SPLEEN,
CC AND PRONEPHROS. VERY LOW EXPRESSION IN LIVER. NO EXPRESSION IN
CC TESTIS, BRAIN, EYE AND MUSCLE.
CC -!- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AT DAY 3-4 IN THE YOLK SAC
```

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 11:46:20 ; Search time 13.55 Seconds
(without alignments)
436.020 Million cell updates/sec

Title: US-09-019-348-8
Perfect score: 126
Sequence: 1 ERDENLKGPMGNAEPEI.....VLDRLASNAVKRSMPOKRF 209

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	684.5	60.8	519	1 IKAR_HUMAN	Q13422 homo sapien
2	676.5	60.1	518	1 IKAR_CHICK	Q42410 gallus gall
3	675	59.9	522	1 IKAR_ONCMY	Q13089 oncorhynchu
4	662.5	58.8	517	1 IKAR_MOUSE	Q31267 mus musculu
5	618	54.9	526	1 HELI_MOUSE	P81183 mus musculu
6	317.5	28.2	1350	1 XFIN_XENLA	P08045 xenopus lae
7	299	26.6	803	1 ZN43_HUMAN	P28160 homo sapien
8	297	26.4	485	1 ZN42_HUMAN	P28698 homo sapien
9	296	26.3	1191	1 ZN91_HUMAN	Q05481 homo sapien
10	294	26.1	553	1 ZF37_MOUSE	P17141 mus musculu
11	293	26.0	572	1 KID1_MOUSE	O61751 mus musculu
12	290	25.8	347	1 KR2_MOUSE	P08043 mus musculu
13	289	25.7	292	1 OZF_BOVIN	Q28131 bos taurus
14	289	25.7	474	1 Z141_HUMAN	Q15938 homo sapien
15	288	25.6	292	1 OZF_HUMAN	Q15072 homo sapien
16	288	25.6	379	1 Z11B_HUMAN	Q06732 homo sapien
17	288	25.6	629	1 Z195_HUMAN	O14628 homo sapien
18	288	25.6	595	1 ZN85_HUMAN	Q03923 homo sapien
19	287.5	25.5	439	1 Z028_XENLA	P18747 xenopus lae
20	285.5	25.4	614	1 ZF29_MOUSE	Q07230 mus musculu
21	285	25.3	506	1 Z157_HUMAN	P51786 homo sapien
22	285	25.3	848	1 Z33A_HUMAN	Q06730 homo sapien
23	284.5	25.3	626	1 Z189_HUMAN	O75820 homo sapien
24	282.5	25.1	510	1 CF2_DROME	P20385 drosophila
25	282.5	25.1	580	1 ZF35_MOUSE	P15620 mus musculu
26	282.5	25.1	428	1 ZN83_HUMAN	P51522 homo sapien
27	282	25.0	393	1 MLZ4_MOUSE	Q03309 mus musculu
28	282	25.0	336	1 ZG57_XENLA	P18729 xenopus lae
29	281.5	25.0	469	1 Z135_HUMAN	P52742 homo sapien
30	280.5	24.9	589	1 Z132_HUMAN	P52740 homo sapien
31	280	24.9	321	1 Z177_HUMAN	Q13360 homo sapien
32	280	24.9	543	1 ZN08_HUMAN	P17098 homo sapien
33	279	24.8	576	1 KID1_RAT	Q02975 rattus norv
34	279	24.8	247	1 Z020_XENLA	P18744 xenopus lae
35	278.5	24.7	428	1 ZF26_MOUSE	P10076 mus musculu
36	278.5	24.7	229	1 ZN22_HUMAN	P17026 homo sapien
37	278	24.7	636	1 ZF90_MOUSE	Q61967 mus musculu
38	278	24.7	682	1 ZN45_HUMAN	Q02386 homo sapien
39	277	24.6	519	1 ZN35_HUMAN	P13682 homo sapien
40	276.5	24.6	726	1 Z184_HUMAN	O99676 homo sapien
41	276.5	24.6	242	1 Z37A_HUMAN	P17032 homo sapien
42	275	24.4	453	1 Z06_XENLA	P18749 xenopus lae
43	274	24.3	686	1 ZN07_HUMAN	P17097 homo sapien

ALIGNMENTS

```
RESULT 1
IKAR_HUMAN
ID IKAR_HUMAN STANDARD; PRT; 519 AA.
AC Q13422; 000598;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1).
GN IKAROS OR IK1 OR OR LYF1.
OS HOMO SAPIENS (HUMAN)
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX TISSUE=BONE MARROW;
RC MEDLINE; 96252222.
RA NIETFIELD W., MEYERHANS A.;
RT "Cloning and sequencing of hik-1, a cDNA encoding a human homologue
of mouse Ikaro/Lyf-1."
RL IMMUNOL. LETT. 49:139-141(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96132884.
RA MOLNAR A., WU P., LARGESPADA D.A., VORTKAMP A., SCHERER S.,
COPELAND N.G., JENKINS N.A., BRUNS G., GEORGIOPOULOS K.;
RT "The Ikaro gene encodes a family of lymphocyte-restricted zinc
finger DNA binding proteins, highly conserved in human and mouse."
J IMMUNOL. 156:585-592(1996).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE).
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED.
CC DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMUS, SPLEEN AND
PERIPHERAL BLOOD LEUCOCYTES AND LYMPH NODES. LOWER EXPRESSION IN
BONE MARROW AND SMALL INTESTINE.
CC -!- ALTERNATIVE PRODUCTS: SIX FORMS OF THE PROTEIN: ISOFORMS IK1-IK6;
ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS
THAT OF IK1.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U40462; G1289371; -
CC EMBL; S80876; G1911483; -
CC DR DRBL; 603023; -
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
CC DR PFAM; PF00096; zf-C2H2; 5.
CC KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
FT DOMAIN 117 224 ZINC-FINGERS I.
FT ZN_FING 117 139 ZINC-FINGERS II.
FT ZN_FING 145 167 C2H2-TYPE.
FT ZN_FING 173 195 C2H2-TYPE.
FT ZN_FING 201 224 C2H2-TYPE.
FT ZN_FING 462 484 C2H2-TYPE.
FT ZN_FING 490 514 C2H2-TYPE.
```

44 273.5 24.3 169 1 MFGL_MOUSE p16372 mus musculu
45 273 24.2 407 1 OZF_RAT Q62981 rattus norv

Db 249 QKTHESKPOODSPNCEKTFEOTESAP-KTATNDQLHESA--GLEKVPPELPEATNSV 300

Search completed: November 6, 1999, 08:08:52
Job time: 72 sec

Qy	24	YSEYENYENIKULERHVVSFDSRPTSGKKNCDVCG-----LSCI-----	64
		! : : : ! ! : ! ! : ! ! : ! ! :	
Db	151	YVKVAFKFSN--SNRHEIRHTKKK-----FKCTKCGKSGFMISCLTEHSRIHTRVNFYK	204
Qy	64	-----SFN-----VLMVHKRSHTGERPQCNOCGASFQKGNLLRHILKLTCEKPFKCHLCN	115
		! ! : ! ! ! ! ! ! ! ! : ! ! : ! ! : ! ! : ! ! : ! ! :	
Db	205	EECGKAFNWSSTLTTKHKRIHTGKPKYCECGKAFNQSLLKHKHIXHTGKPKYCEEG	264
Qy	116	-----YACQ-----RRDALGHLRHSVEKPKYCFCGRSYK	147
		! : ! ! : ! ! ! ! ! ! ! ! : ! ! : ! ! : ! ! : ! ! :	
Db	265	KAFNRSTLTTHKIIHTGEKPKYCECGKAFNRSSLTTHRKHTGKPKYCECGKAFK	324
Qy	148	QRSLEEHK	156
		! : ! ! !	
Db	325	OSSNLTTHK	333

RESULT 15
S06556
finger protein (clone XlcoF28) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990.#text_change 31-Dec-1993
C:Accession: S06556
R.J. Mielkef, W.: El-Baradi, T.; Mentzel, H.; Pieker, T.; Koester, M.; Poeting, A.; Knoechel, J.
J. Mol. Biol. 208, 639-659, 1989
A:Title: Second-order repeats in Xenopus laevis finger proteins.
A:Reference number: S05632; MUID:90040698
A:Accession: S06556
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-439 <NIE>
C:Keywords: DNA binding; zinc finger

```

Query Match      25.5%; Score 287.5; DB 2; Length 439;
Best Local Similarity 36.6%; pred. No. 6.6e-18;
Matches 64; Conservative 28; Mismatches 54; Indels 19; Gaps 5;

QY      36 LRRVVVSFDSRPTSGKMNCDVCGLSGTSFNLVLMVKRSRHTGERPFCNOCGASFTQGN 95
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      133 LVEHQRHPTGDRP-----FOCTECKSFKSKSLVAVHLRIHTGEKPYKCECDKSFQSOST 188
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      96 LRRHKLHTGKPKFKLCLNCAQRRDALTHSVKPKYKCFGCRSYKQRSSLEEH 155
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      189 LVVHQRHHTGERPQCSCHEKSFYHYAFVVHRTHTGEKPYKSCMDKAYSQRSNLKLH 248
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      156 KE-----RC-RTFLQSTDPGDTASAEARHIKAEMGSEKALVLDRLASNV 198
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

A:Reference number: A90892; MUID:87131089
A:Accession: B26595
A:Molecule type: DNA
A:Residues: 75-347 <CH2>
A:Cross-references: GB:M15709; NID:g19335
C:Genetics:
A:Gene: mkx2
C:Keywords: DNA binding; zinc finger

Query Match	25.8%	Score 290;	DB 2;	Length 347;
Best Local Similarity	34.8%;	Pred. NO. 3.1e-18;		
Matches 54;	Conservative 34;	Mismatches 61;	Indels 6;	Gaps 2;

QY	2	RDENVLKSEPMGNAEPEIPYSRYSENYENTKLRHVVSFDSRPTSGKMNCDCVGLS	61
Db	111	QSSLLKQRIHTGKPYTCNVCKHIFERSLTV--HQRTHTGKPE---YKCEHGKRA	164
OY	62	CIGFNVLMVHKRSHGTGERPFCQNOCGASTQGNLLRLHKLTGKPFKCHLCNTACQRR	121
Db	165	FSOSMNLTVHQRTHTEKPYQCEKCAFKNSLSIQHERIHTGKPYKCHDCGKAFSKN	224
QY	122	DALTGLRTHSVKPKYKCFCKGSRYSKORSLEHK	156
Db	225	SSLTQRRRIHTGKPYECMCKGKHFTGRSSLTVHO	259

RESULT 11
I54338
zinc finger protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Aug-1997
C:Accession: I54338
R:Tommerup, N.; Aggaard, L.; Lund, C.L.; Boel, E.; Baxendale, S.; Bates, G.P.; Lehrach,
Hum. Mol. Genet. 2, 1571-1575, 1993
A:title: A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate gene for the w
A:Reference number: I54338; MID:94093541
A:Accession: I54338
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-474 <RES>
A:Cross-references: GB:I15309; NID:g347905; PID:g347906
C:Genetics:
A:Gene: GDB:ZNF141
A:Cross-references: GDB:i37046; OMIM:194648
A:Map position: 4p16.3-4p16.3

Query Match 25.7%; Score 289; DB 2; Length 474;
Best Local Similarity 34.6%; Pred. No. 5.2e-18;
Matches 56; Conservative 22; Mismatches 60; Indels 24; Gaps 2;

Oy	16	EPEEIPYSREYNEVNI-----KLRHVVSFSSRPTSGKWNC	55
		: : : :	
Db	202	EECGAKFKWSLIFNEHKRIHTCEKPTECEGSIETTSFHAKHIIITGEKP----	257
Oy	56	DVGCLSCISFNVLWVKHRSHTERPPQCNQGASSETQGNLLRLHKIHLKTGETKPFKCHLCN	115
		: : : : : :	
Db	258	EECGAFNRFTLTUKKRITHAGEKPITCECKRIETSSNFAGKHRIHTGEKPYKECEGG	317

Qy 116 YACORRALTGHRLTHSVKPYKCEGSRYSQRSSLEEHKE 157
 318 KAFNRSTTLTKHRIHTGEKPYTCECGKAFRQSKLNEHKK 359
 Db

RESULT 12
S43826
finger protein OZF, krueppel-related - human
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-1999
C:Accession: S43826
R:Le Chalony, C.; Prosperi, M.T.; Haluza, R.; Apiou, F.; Dutrillaux, B.; Goubin, G.
J. Mol. Biol. 236, 399-404, 1994

A:Title: The OZF gene encodes a protein consisting essentially of zinc finger motifs.
A:Reference number: S43826; MUID:94149744
A:Accession: S43826
A:Molecule type: mRNA
A:Residues: 1-292 <LEC>
A:Cross-references: EMBL:X70394; NID:g468707; PID:g468708
C:Genetics:
A:Gene: GDB:ZNF146; OZF
A:Cross-references: GDB:230289; OMIM:601505
A:Map position: 19q13.1-19q13.1
C:Keywords: DNA binding; zinc finger

Query Match	25.6%	Score 288;	DB 2;	Length 292;
Best Local Similarity	30.4%	Pred. No. 3.9e-18;		
Matches 62; Conservative	44;	Mismatches 68;	Indels 30;	Gaps 5;

```

Qy 4 ENVLKSEPMGNAEPEIPYSYREYNEYENIKLERHVVSFDSRP-----TSGKN 54
    ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 85 ENLLTHOKIHTGKEPFECKDCGKAFIQKS--LIRHQRTHTGKEPVCCKGTSGKN 142
    ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Qy 55 -----CDVCGLSGISENVLMVHKRSHTGSRPFCNOCGASFQKGNLLRH 99
    ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 143 LTHEKTHIGKEPFCSECGTAFQKKYLIKQNIHTGKEPCECNCGKAFSORTSLIVH 202
    ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Qy 100 IKLHTGKPKFCHLCNACORRDALGHLRTHSVEPKYCFEGRSYKQRSSLEHKRC 159
    ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 203 VRIHSGDKPYECNVCCKAFSQSSSLIVHVRSHTGEPYCNCGKAFSFTLALHL-RI 261
    ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Qy 160 RTF---LQSTDPDGDTASAEARHIK 180
    | : : | : : | : : | : : | : : | : : | : : | : :
Db 262 HTGKKPYQCSECGKAFSOKSHHIR 285
    | : : | : : | : : | : : | : : | : : | : : | : :

```

```

RESULT 13
S33994
finger protein ZNF11b - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_
C:Accession: S33989; S33994
R:Tunnacilife, A.; Liu, L.; Moore, J.K.; Leversha, M.A.;
Nucleic Acids Res. 21, 1409-1417, 1993
A:Title: Duplicated KOX zinc finger gene clusters flank t
A:Reference number: S30238; MUID:9319119
A:Accession: S33989
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 65-287 <TU2>
A:Cross-references: EMBL:X68685; NID:g296453; PID:g938236
A:Accession: S33994
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <TUN>
A:Cross-references: EMBL:X68684
C:Genetics:
A:Gene: GDB:ZNF11b
A:Cross-references: GDB:128306; OMIM:194522
A:Map position: 10q11.2-10q11.2
C:Keywords: zinc finger

```

Query Match 25.68; Score 288; DB 2; Length 378;
Best Local Similarity 49.08; Pred. No. 5.1e-18;
Matches 50; Conservative 20; Mismatches 32; Indels 0; Gaps 0;

[illegible]

Db 75 NHCGVLSHKQGLDHTHTGKPYECNECGIAFSQKSHLVLRHTHTGKPYECQCG 134
Qy 116 YACORRDLTGLTHSVKPYKCFGRSKYKQSSLEEH-----KERCRTFL 163
Db 135 KAHGKHLDHTLRIHTGKPYKNECGKTPRHSNLMQHLRSHTGKPYECKGKSR 194
Qy 164 QSTDPGDTASAEARHKAEMG 184
Db 195 YN-----SSLTEHVRHTG 208
RESULT 7
A:0751
finger protein MZF1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 24-Sep-1998
C:Accession: A40751
R:Hromas, R.; Collins, S.J.; Hickstein, D.; Raskind, W.; Deaven, L.L.; O'Hara, P.; Hagen
J. Biol. Chem. 266, 14183-14187, 1991
A:Title: A retinoic acid-responsive human zinc finger gene, MZF-1, preferentially expres
A:Reference number: A40751; MUID:913117761
A:Accession: A40751
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <HRO>
A:Cross-references: GB:M58297; NID:g189043; PID:g189044
C:Genetics:
A:Gene: GDB:ZNF42; MZF-1
A:Cross-references: GDB:125898; OMIM:194550
A:Map position: 19q13.2-19q13.4
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 26.4%; Score 297; DB 2; Length 485;
Best Local Similarity 41.0%; Pred. No. 1.1e-18;
Matches 57; Conservative 22; Mismatches 50; Indels 10; Gaps 2;

Qy 18 PEIPYSYSREYENIKLERHVVSDFSSRPTSGKMNCDVGLSCISFNVLVHVKRSHTG 77
Db 351 PCGGRFS-----QRLALTRHQRHTGKPY-----YHCGEGLGFTQVSRUTEHQRIHTG 400
Qy 78 EPPFCNCGASFTQKGNLRLHKLHTGKPKPKCHLCNYACORRDLTGLTHSVKPY 137
Db 401 ERPFACPECGSGFRGHANLTHRRHTGKPYACPECGKAFQRPTLTQHLRTHRRKPF 460
Qy 138 KCEFCGRSKYKQSSLEEH 156
Db 461 ACQDCGRRFHQSTKLIQHQ 479

RESULT 8
S35305
zinc finger protein ZNF91 - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: S35305
R:Bellegroid, E.J.; Marine, J.C.; Ried, T.; Lecocq, P.J.; Riviere, M.; Amemiya, C.; Pond
EMBO J. 12, 1363-1374, 1993
A:Title: Clustered organization of homologous KRAB zinc-finger genes with enhanced expre
A:Reference number: S35305; MUID:93223677
A:Accession: S35305
A:Molecule type: mRNA
A:Residues: 1-1191 <BEL>
A:Cross-references: EMBL:L11672; NID:g186773; PID:g186774
A:Note: The authors translated the codon GCA for residue 750 as Thr and GCT for residue
C:Genetics:
A:Gene: GDB:ZNF91; HPF7; HTF10
A:Cross-references: GDB:132284
A:Map position: 19p12-19p12
C:Keywords: DNA binding; zinc finger

Query Match 26.3%; Score 296; DB 2; Length 1191;

Best Local Similarity 34.9%; Pred. No. 3.3e-18;
Matches 66; Conservative 29; Mismatches 78; Indels 16; Gaps 4;
Qy 12 MGNAPPEIPYSYSREYENIKLERHVVSDFSSRPTSGKMNCDVGLSCISFNVLVH 71
Db 1015 MHTGKPYKCECGKAFN--RSSKLTHTKIHTGKPY-----YKCECGKAFISSSTLNGH 1068
Qy 72 KRSHTGERPFQCNCGASFTQKGNLRLHKLHTGKPKPKCHLCNYACORRDLTGLTHRTH 131
Db 1069 KRHTREKPYKCECGKAFSOSTLTNRKRLHTGKPKYKCECGKAFKSSALTKKHIIH 1128
Qy 132 SVKPYKCFGRSKYKQSSLEEHKERCRTFLQSTDPGDTASAEARHKAEMGSERLVL 191
Db 1129 TGEKPYKCECGKAFNOSILTNRKK-----IHTITVPIPLLWE-----AEAGSGRGQEM 1178
Qy 192 DRLASNVAK 200
Db 1179 ETILANTVK 1187
RESULT 9
S22954
finger protein zfp-37 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
C:Accession: S22954
R:Burke, P.S.; Wolgemuth, D.J.
Nucleic Acids Res. 20, 2827-2834, 1992
A:Title: zfp-37, a new murine zinc finger encoding gene, is expressed in a developmen
A:Reference number: S22954; MUID:92310982
A:Accession: S22954
A:Molecule type: mRNA
A:Residues: 1-553 <BUR>
A:Cross-references: EMBL:X64413; NID:g55474; PID:g55475
C:Superfamily: zinc finger protein zfp-36
C:Keywords: DNA binding

Query Match 26.1%; Score 294; DB 2; Length 553;
Best Local Similarity 38.0%; Pred. No. 2.2e-18;
Matches 57; Conservative 29; Mismatches 44; Indels 20; Gaps 3;

Qy 48 PTSK-MNCDVGLSCISFNVLVHVKRSHTGERPFQCNCGASFTQKGNLRLHKLHTGE 106
Db 208 PGSGKPYECNCHGKVLSHKQGLDHTHTGKPYKCECGKAFKSSALTKKHIIH 267
Qy 107 KPEKCHLCNYACORRDLTGLTHSVKPYKCFGRSKYKQSSLEEH----- 156
Db 268 KPYECGCGKAHGHKHALTDHLRIHTGKPKYKNECGKTPRHSNLMQHLRSHTGKPYE 327
Qy 156 -KERCRTFLQSTDPGDTASAEARHKAEMG 184
Db 328 CKEGKSFYRN-----SSLTEHVRHTG 350

RESULT 10
S00549
developmental control protein Kr2 - mouse
N:Alternate names: mkr2
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 20-Mar-1998
C:Accession: S00549; B26595
R:Chowdhury, K.; Dressler, G.; Breier, G.; Deutsch, U.; Gruss, P.
EMBO J. 7, 1345-1353, 1988
A:Title: The primary structure of the murine multifinger gene mkr2 and its specific e
A:Reference number: S00549; MUID:88312581
A:Accession: S00549
A:Molecule type: mRNA
A:Residues: 1-347 <CHO>
A:Cross-references: EMBL:Y00850; NID:g52810; PID:g52811
R:Chowdhury, K.; Deutsch, U.; Gruss, P.
Cell 48, 771-778, 1987
A:Title: A multigene family encoding several "finger" structures is present and diffe

Db 114 HKCGYGRSYKQRSSLEEHKERCNHYLESMGLPGVCPVKEETNHNEMAEADLCKIGAERS 173

QY 189 LVLDRLASNVAKRKSSMPQKF 209
 |
Db 174 LVLDRLASNVAKRKSSMPQKF 194

RESULT 3

B56229
lymphoid transcription factor Ikaros/LyF-1, form V - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995
C:Accession: B56229
R:Haehn, K.; Ernst, P.; Lo, K.; Kim, G.S.; Turck, C.; Smale, S.T.
Mol. Cell. Biol. 14, 7111-7123, 1994
A:Title: The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced transcripts
A:Reference number: A56229; MUID:95021239
A:Accession: B56229
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-427 <HAH>
A:Cross-references: GB:S74518
C:Keywords: alternative splicing

Query Match 50.1%; Score 564; DB 2; Length 427;
Best Local Similarity 75.9%; Pred. No. 4.6e-42;
Matches 107; Conservative 11; Mismatches 15; Indels 8; Gaps 2;

77	GERPFCNQCGASFTQKGNLLRHLYKLHTGKEPKFKCHLCNYACORRDALTGHLETHSVEXP	136
	:	
Db	54 GERPFCNQCGASFTQKGNLLRHLYKLHSEKPFKFKCHLCNYACRRRDLTGHLETHSVGXP	113

```
QY      137 YKCFGGSYKQRSLEEHKERCTFLQSTD-PG-----DTASAEARHIKAEMGSEFA 188
```

Db	114	HKCGYCGRSYKQRRSLEEKKERCHNYLESGLPGMPVYKEETNNHMAEDLCKIGAERS	173
Qy	189	LVLDRLASNVAKRRKSSMPQKF	209

Db 174 LVLDRLASNVAKRKSSMPQKF 194

RESULT 4

S00647
 finger protein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Seq-Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 31-Dec-1993
 C:Accession: S00647
 R:Ruiz i, Altaba, A.; Perry-O'Keeffe, H.; Melton, D.A.
 EMBIO J. 6, 3065-3070, 1987
 A:Title: Xfin: an embryonic gene encoding a multifingered protein in Xenopus.
 A:Reference number: S00647; UID:88082679
 A:Accession: S00647
 A:Molecule type: mRNA
 A:Residues: 1-1350 <RUI>
 A:Cross-references: EMBL:X06021
 A:Note: it is uncertain whether Met-1, Met-11, Met-16, Met-38, or Met-39 is the initiator
 C:Keywords: DNA binding; zinc finger

Query Match	28.2%	Score	317.5	DB 2:	Length	1350;	
Best Local Similarity	36.9%	Pred. No.	4.9e-20;				
Matches	56:	Conservative	33;	Mismatches	59;	Indels	21;

```
004 ENVLKSEPMGNAEPEIPYSYSREYNEYENIKLERHVVSFDSRPTSGMKN--CDVCGL 60
```

```

308 DSWGTDPLSSQNVASSPYSCSKRTKRWK-----SFLNHQOHTSREKPYLCSHCNK 361
202 61 SCISFNVLMVHKRSHTGTPQCNOCGASTQKGNLLRHLKIHTGEKPFKCHLGNVACQR 120
      | : | : |:| |:| |:| | : | : |:| |:| |:| |:| |:| |:| |:| |:| |:|

```

Db 362 GFIONSDLVHRFTHTERPYOACECHKGFIOKSLDKVHLRHTHTEGKPFKCSHCCKFTE 421

QY 121 RDALTGHLRTHSVEKYKCEFCGRYSQRSSLEEHKE-----RC-----RTFLQSTD 167
| | | | | | | | | | | | | | | : | | | | : |
| | | | | | | | | | | | | | | : | | | | : |

Db 422 RSALKRHQTHTEGKYPKSCDCKEKTQRSNLLHORHTHGERPYKCTLCDRTFIQNSD 480

RESULT 5

S26823
finger protein ZNF43 - human
C:Species: Homo sapiens (man)
C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1998
C:Accession: S26823
R:Loebering, R.; Trowsdale, J.
Nucleic Acids Res. 19, 2921-2928, 1991
A:title: A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell
A:reference number: S26823; MUID:91279444
A:Accession: S26823
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-803 <LOV>
A:Cross-references: EMBL:X59244; NID:g38031; PID:g38032
C:Genetics:
A:Gene: GDB:ZNF43; HTF6
A:Cross-references: GDB:128653
A:Map position: 19p13.1-19p12
C:Keywords: DNA binding

Query Match	26.6%	Score 299;	DB 2;	Length 803;
Best Local Similarity	35.3%;	Pred. No. 1.2e-18;		
Matches 61;	Conservative 26;	Mismatches 56;	Indels 30;	Gaps 3;

QY 4 ENVLKSEPMCAEEPEIPYSREYNEYENI-----KLERHVVSF 43

Db 473 EKP YKCECGKA-----FSRSSNLTKHKIHI EKKPYKCECGKAFKWSKLTETHKITH 526

Qy 44 DSRPTSGRMNCDCVGLSCISFNVLVMVHKRSHTGERPQCNCQSGASFTQGNLLRIKLH 103

52/ TGEKP----YKCECGKAFNHFSILTKHKRIHTGEXPKCECGKAFQSSNLTTHK
D0
QY 104 TGEKPFKCHLCNYACQRRDALTGHLRTHSVEKPKCFECGRSYKQPSLSEHK 1566

Db 583 TGEKFYCEECGKAFQTSSNLTTHKKIHTGGKPYKCEECGKAFNQFSTLTKKH 635

RESULT 6

S10245
finger protein, testis - mouse
C:Species: Mus musculus (house mouse)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 10-Sep-1997
C:Accession: S10245
R:Nelki, D.; Dudley, K.; Cunningham, P.; Akhavan, M.
Nucleic Acids Res. 18, 3655, 1990
A:Title: Cloning and sequencing of a zinc finger cDNA expressed in mouse testis.
A:Reference number: S10245; MUID:90301500
A:Accession: S10245
A:Molecule type: mRNA
A:Residues: 1-411 <NEL>
A:Cross-references: EMBL:X52533; NID:g53456; PID:g53457
C:Keywords: DNA binding; zinc finger

Query Match
Best Local Similarity 26.5%; Score 298.5; DB 2; Length 411;
32.8%; Pred. No. 6.7e-19;

Matches	66;	Conservative	34;	Mismatches	66;	Indels	35;	Gaps
Oy	6	VLKSEPMGNAPEEIPYSREYNEY-----	ENIKLERHVVFDSRPTSGK-MNC	55				

```
Db      | |:| | | : | |:| | | | | | | |
21 VTtKpAKNVRK-----YARHSASHTKEDIQTGEKRKSHCHRTSPKPEKAPGSGKPYEC 74
```

QY 56 DYCGLSGISFNVLWVHKRSHRTGERPFQCNOCCGASFTQKGNLLRHKLHTGEKPPFKCHLCLN 115
 : | : | : | : | : | : | : | : | : | : | : | : | : | : |

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 08:08:51 ; Search time 18.18 Seconds
(without alignments)
460.599 Million cell updates/sec

Title: US-09-019-348-8
Perfect score: 1126
Sequence: 1 ERDENVKSEPMGNAEPEI.....VLDRLASNVAKRKSSMPQKF 209

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR_60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	671	59.6	518	2 A56355	DNA-binding protel
2	564	50.1	431	2 I59572	Ikaros DNA binding
3	564	50.1	427	2 B56329	lymphoid transcrip
4	317.5	28.2	1350	2 S00647	finger protein - A
5	299	26.6	803	2 S26823	finger protein ZNF
6	298.5	26.5	411	2 S10245	finger protein, te
7	297	26.4	485	2 A40751	finger protein MZF
8	296	26.3	1191	2 S35305	zinc finger protel
9	294	26.1	553	2 S22954	finger protein zfp
10	290	25.8	347	2 S00549	developmental cont
11	289	25.7	474	2 I54338	zinc finger protel
12	288	25.6	292	2 S43826	finger protein O2F
13	288	25.6	378	2 S33894	finger protein ZNF
14	288	25.6	595	2 G02075	transcription repr
15	287.5	25.5	439	2 S06556	finger protein (cl
16	285.5	25.4	614	2 JH0500	zinc finger protel
17	284	25.2	732	2 S47073	finger protein HZF
18	282.5	25.1	510	2 A36901	chorion transcript
19	282.5	25.1	428	2 A32891	finger protein 1,
20	282.5	25.1	580	2 A37107	spermatogenesis pr
21	282	25.0	336	2 S06578	finger protein (cl
22	282	25.0	393	2 JN0533	finger protein pML
23	281.5	25.0	427	2 A35659	krueppel-related p
24	281.5	25.0	469	2 I38600	zinc finger protel
25	280.5	24.9	589	2 I38598	zinc finger protel
26	280	24.9	543	2 B34612	zinc finger protel
27	279	24.8	247	2 S06553	finger protein (cl
28	279	24.8	576	2 A48157	renal transcriptio
29	278.5	24.7	229	2 A48927	kruppel-like zinc
30	278.5	24.7	428	2 S03677	finger protein (cl
31	278	24.7	247	2 S47070	finger protein HZF
32	278	24.7	636	2 I48689	transcription acti
33	277	24.6	519	2 A38073	gene NK10 protein
34	277	24.6	693	2 I37570	zinc finger protel
35	277	24.6	295	2 A46017	zinc finger protel
36	276.5	24.6	546	2 I49636	DNA-binding protel
37	276.5	24.6	242	2 S30238	finger protein ZNF
38	275.5	24.5	491	2 S00753	finger protein HF
39	275	24.4	453	2 S06544	finger protein (cl

40 275 24.4 591 2 S65088 finger protein XFO
41 274 24.3 686 2 A34612 zinc finger protel
42 273.5 24.3 169 2 A39240 finger protein mfg
43 273 24.2 651 2 B32891 finger protein 2,
44 272.5 24.2 210 2 S06577 finger protein (cl
45 271.5 24.1 462 2 S10397 finger protein kox

ALIGNMENTS

RESULT 1
A56355
DNA-binding protein Ikaros form 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C:Accession: A56355
R: Molnar, A.; Georgopoulos, K.
Mol. Cell. Biol. 14, 8292-8303, 1994
A:Title: The Ikaros gene encodes a family of functionally diverse zinc finger DNA-bin
A:Reference number: A56355; MUID:95059058
A:Accession: A56355
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-518 <MOL>
C:Keywords: alternative splicing; DNA binding

Query Match 59.6%; Score 671; DB 2; Length 518;
Best Local Similarity 73.7%; Pred. No. 2.6e-51;
Matches 126; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

QY 47 RPTSGKMGDCVGLSCISFNVLVHVRSHSGTGERPFCNCGASFTQKGNLLRHKLTGE 106
DB 111 RLPNGKLKDCIGVICGPNVLMVHVRSHSGTGERPFCNCGASFTQKGNLLRHKLTGE 170
QY 107 KPFKHLNVCACRRDALTGHLRTHSVKPYKCFGCRSYKQRSSLEEHKERCRTFLOST 166
DB 171 KPFKHLNVCACRRDALTGHLRTHSVKPYKCFGCRSYKQRSSLEEHKERCRTFLOST 230
QY 167 D-PG-----DTASAEARHKAEMGERALVLDRLASNVAKRKSSMPQKF 209
DB 231 GLGPMYPIKEETNHNHMAEDLCKIGALERSLVLDRLASNVAKRKSSMPQKF 281

RESULT 2
Ikaros DNA binding protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C:Accession: I59572
R: Georgopoulos, K.; Moore, D.D.; Derfler, B.
Science 258, 808-812, 1992
A:Title: Ikaros, an early lymphoid-specific transcription factor and a putative media
A:Reference number: I59572; MUID:93068267
A:Accession: I59572
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-431 <RES>
A:Cross-references: GB:I03547; NID:g198286; PID:g198287
C:Genetics:
A:Gene: Ikaros

Query Match 50.1%; Score 564; DB 2; Length 431;
Best Local Similarity 75.9%; Pred. No. 4.7e-42;
Matches 107; Conservative 11; Mismatches 15; Indels 8; Gaps 2;

QY 77 GERPFQNCQASFTQKGNLLRHKLTGKPKFCHLCNVCACRRDALTGHLRTHSVK 136
DB 54 GERPFQNCQASFTQKGNLLRHKLTGKPKFCHLCNVCACRRDALTGHLRTHSVK 113
QY 137 YKCFGCRSYKQRSSLEEHKERCRTFLOSTD-PG-----DTASAEARHKAEMGSRA 188

MOLECULE TYPE: protein
US-08-398-590A-40

Query Match 20.0%; Score 225; DB 2; Length 676;
Best Local Similarity 27.9%; Pred. No. 2e-17;
Matches 46; Conservative 32; Mismatches 73; Indels 14; Gaps 2;

QY 1 ERDENLVKSEPMGNAPEPIYS-----YGREYNEYENIKLERHVVVSFDSRPTS 50
Db 270 ESAEQAKARESGSSTAEGDSKPIRCDCGYNTNRVDYHTAHLKHHTHTRAGDNER--- 327
QY 51 GRMNCVCLGICISFNLVHVRSHRTGERPFCNOCGASFTQKGNLLRHILKHTGEKPK 110
Db 327 -VYKCIITTYTVSEYHVRKHLRNFPRKYVTCGKCNFYSDRKNVYQHVRTHTGERPYK 385
QY 111 CHLCNYACORRDALTGHLRTHSVKPYKCFGRSYKORSSLEEH 155
Db 386 CELCPYSSOKTHLRHMRTHSGERPFKCDQCSYVASNOHEVTRH 430

RESULT 14
US-08-718-661-2
Sequence 2, Application US/08718661
Patent No. 5876972
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
NUMBER OF SEQUENCES: 15
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
FILING DATE: US/08/718,661
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 667 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-718-661-2

Query Match 19.9%; Score 224; DB 2; Length 667;
Best Local Similarity 31.8%; Pred. No. 2.5e-17;
Matches 50; Conservative 25; Mismatches 70; Indels 12; Gaps 5;

QY 55 CDVCGLSICISFNLVHVRSHRTGERPFCN--OCGASFTQKGNLLRHILKHTGEKPKCH 112
Db 6 CQCKGRSFVTEKFTIHNYSHSRERPFKCAECGKAFVSKYKLMRHMATHSPQKHQCT 65
QY 113 LCNYACORRDALTGHLRTHSVK-PYKCFGRSYKORSSLEEHKERCFTLQSTDPGDT 171
Db 66 HCEKTFNRKDLKLNHLOTHDPNKSISYACDDCGKKYHTMLGYKRH-----LALHSASNGDL 120
QY 172 ASAEARHIKAEMGSEALVLDRLASNAVKRKSMPQK 208
Db 121 TCGVC---TLELGSFVL-LDHLKSHAEEKANQAPRE 153

RESULT 15
US-08-224-482-4
Sequence 4, Application US/08224482
Patent No. 5837692
GENERAL INFORMATION:
APPLICANT: Mercola, Dan
APPLICANT: Adamson, Eileen D.
TITLE OF INVENTION: Inhibition of the Mitogenic Activity of

TITLE OF INVENTION: PDGF by Mammalian EGR
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 07-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-ME 9913
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 543 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-224-482-4

Query Match 18.7%; Score 210.5; DB 2; Length 543;
Best Local Similarity 36.4%; Pred. No. 6.6e-16;
Matches 40; Conservative 19; Mismatches 46; Indels 5; Gaps 2;

QY 78 ERPFQC--NOCGASFTQKGNLLRHILKHTGEKPKFCHLCNYACORRDALTGHLRTHSVK 135
Db 335 ERPYACPVESCDRRFSRSDLTIRIHTGQKPFQCRICMNFERSDHLTTHRTHTGCK 394
QY 136 PYKCFGRSYKORSSLEEHKERCFTLQSTDPGDTASAEARHIKAEMGS 185
Db 395 PFACDICGRKFARS---DERKRHTKIHLRQDKKADKSVVASSATSSLS 441

Search completed: November 6, 1999, 09:41:36
Job time: 68 sec

```

; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60897/RFT/RMS
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
;
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 676 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```


TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-820-170A-10

Query Match 23.9%; Score 269; DB 2; Length 711;
Best Local Similarity 40.7%; Pred. No. 2e-22;
Matches 48; Conservative 22; Mismatches 48; Indels 0; Gaps 0;
QY 39 HVVSDSSRPTSGKMNCDVGLSCISFNVLVHKRSHGTGERPFCQNCQSGASFTQKGNLLR 98
DB 461 HLIVHQRSHGTGKPYOCHNCGKSFISKSLDIHRIHTGKPYECSDCGKTFQKSHLNI 520
QY 99 HIKLHTGEKPFKCHLNYACQRDALTGHLRTHSVKPYKCEFCGSRYSKORSLEEKK 156
DB 521 HQIHTGERHVGSECGKAFNQKSIILSMHQRIHTGKPYKCEGKAFYKSKOFKEHQ 578

RESULT 7
US-08-933-750C-17
Sequence 17, Application US/08933750C
Patent No. 5932442
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: BRSTNOT03
CLONE: 641127
US-08-933-750C-17

Query Match 23.4%; Score 263; DB 2; Length 488;
Best Local Similarity 28.1%; Pred. No. 5.6e-22;
Matches 62; Conservative 41; Mismatches 72; Indels 46; Gaps 7;
QY 6 VLKSPMGNAEPEIPYSYREY-----NEYENIKLERHVVSFDSRPTSGKMNCDVC 58
DB 187 ILNTSPDGHPIEK--IHTAKQYEGSCQCKNFSSQSELLLHORDHTEKP-----YKCEOC 240
QY 59 GUSCISFNVLVHKRSHGTGERPFCQNCQSGASFTQKGNLLRHLIKHTGKPKCHLNYAC 118
DB 241 GKGFTRSSLLIHOAVHTDEKPYKDKCGKGTFRSSLLIHHAVHTGKPKYKCDKCGKGF 300
QY 119 QRDLALTGHLRTHSVKPYKCEFCGSRYSKORSLEEKK-----RC-----RTFLQST 166
DB 301 SQSSKLHIHQHVHTGKPYECECGMSFSQRSNLHIHQHVHTGKPYKCGCGKGFSSQS 360
QY 167 D-----PGDTASAEAR-HIKAEMGSE 186
DB 361 NLHIHRCIHTGKPYQCYECGKGFSSQSSDLRIHLRVHTGK 401

RESULT 8
US-08-074-967-2
Sequence 2, Application US/08074967
Patent No. 5641672
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, R.S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/074,967
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 706 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-074-967-2

Query Match 22.9%; Score 258; DB 1; Length 706;
Best Local Similarity 38.7%; Pred. No. 3.6e-21;
Matches 48; Conservative 18; Mismatches 54; Indels 4; Gaps 1;

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933.750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 338 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMCINOT01
CLONE: 9337
US-08-933-750C-4

Query Match 24.3%; Score 273.5; DB 2; Length 338;
Best Local Similarity 37.5%; Pred. No. 2.1e-23;
Matches 57; Conservative 22; Mismatches 64; Indels 9; Gaps 2;

Qy 55 CDVCGLSGISFNVLVHVKRSHGTGPFQCNOCGASFTQKGNLLRIKLTGCKPKCHLC 114
Db 194 CSECKIFRNFYFVHKRIHTGEPYVCQCGKGFVSSSLTQHRVHSGRPFECQEC 253
Qy 115 NYACORRDALTGHLRTHSVKPYKCFGCRSYKQRSSLEEHKERCRTFLQSTDGDTASA 174
Db 254 GRTFNDRSAISOHLRTHGTAKPYKQDCGKAFRQSSHLIRH-----QRTHTGERPYA 305
Qy 175 EARHTKAMGSEARALVLDRLASNAVKRSSMP 206
Db 306 CNKCGKAFQTSSH-LIGHQRTHNRTKRRKKOP 336

RESULT 5
US-08-933-750C-14
Sequence 14, Application US/08933750C
Patent No. 5932442
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933.750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: MMLR3DT01
CLONE: 569710
US-08-933-750C-14

Query Match 24.2%; Score 272; DB 2; Length 368;
Best Local Similarity 35.8%; Pred. No. 3.5e-23;
Matches 60; Conservative 29; Mismatches 62; Indels 18; Gaps 5;

Qy 2 RDN-----VLKSEPMGNAEPEIPYSR-----EYNE--YENIKLERHVVSFDSRPTS 50
Db 194 RTENGALAPKQELPSALSHSEVPTLSMGVPIFKYGETCEPKGRFR-----KRNP 247
Qy 51 GKMN-CDVCGLSGISFNVLVHVKRSHGTGPFQCNOCGASFTQKGNLLRIKLTGCKPKF 109
Db 248 KKQICDECGKHFSQGSALILHQRHSGKPYGCVGCKAFSRSSILVQHQVHTGKPY 307
Qy 110 KCHLCNVACORRDALTGHLRTHSVKPYKCFGCRSYKQRSSLEEHKER 158
Db 308 KCLECGKAFSONGLINHQRHTGKPYECVQCCKGKSYSSQSNLFRHQR 356

RESULT 6
US-08-820-170A-10
Sequence 10, Application US/08820170A
Patent No. 5831058
GENERAL INFORMATION:
APPLICANT: Tsutomu, FUJIWARA
APPLICANT: Takeshi, WATANABE
APPLICANT: Masato, HORIE
APPLICANT: Toyomasa, KATAGIRI
TITLE OF INVENTION: HUMAN GENE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/820,170A
FILING DATE:
CLASSIFICATION: 536
TELECOMMUNICATION INFORMATION:

; MEDIUM TYPE: Diskette

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	679.5	60.3	470	2	US-08-465-590-153	Sequence 153, Appl
2	679.5	60.3	470	3	PCT-US95-09345-13	Sequence 13, Appl
3	679	60.3	568	3	PCT-US93-08743-5	Sequence 5, Appl
4	273.5	24.3	338	2	US-08-933-750C-4	Sequence 4, Appl
5	272	24.2	368	2	US-08-933-750C-14	Sequence 14, Appl
6	269	23.9	711	2	US-08-820-170A-10	Sequence 10, Appl
7	263	23.4	488	2	US-08-933-750C-17	Sequence 17, Appl
8	258	22.9	706	1	US-08-074-967-2	Sequence 2, Appl
9	258	22.9	706	2	US-08-933-750C-10	Sequence 10, Appl
10	258	22.9	706	3	PCT-US95-541B-2	Sequence 2, Appl
11	252.5	22.4	728	3	PCT-US94-06669-2	Sequence 2, Appl
12	252.5	22.4	727	3	PCT-US95-08439-9	Sequence 9, Appl
13	225	20.0	676	2	US-08-398-590A-40	Sequence 40, Appl
14	224	19.9	667	2	US-08-718-661-2	Sequence 2, Appl
15	210.5	18.7	543	2	US-08-224-482-4	Sequence 4, Appl
16	210	18.7	387	2	US-08-224-483-8	Sequence 8, Appl
17	209.5	18.6	533	2	US-08-040-548-1	Sequence 1, Appl
18	209.5	18.6	533	2	US-08-466-344-1	Sequence 1, Appl
19	209.5	18.6	486	2	US-08-224-483-2	Sequence 2, Appl
20	208	18.5	84	2	US-08-459-568-62	Sequence 62, Appl
21	208	18.5	84	2	US-08-399-411-62	Sequence 62, Appl
22	207.5	18.4	456	2	US-08-040-548-2	Sequence 2, Appl
23	207.5	18.4	456	2	US-08-466-344-2	Sequence 2, Appl
24	207.5	18.4	455	2	US-08-224-483-6	Sequence 6, Appl
25	206	18.3	89	2	US-08-040-548-8	Sequence 8, Appl
26	205	18.3	89	2	US-08-466-344-8	Sequence 8, Appl
27	204.5	18.2	429	1	US-08-234-783-4	Sequence 4, Appl
28	204.5	18.2	429	1	US-08-456-907-4	Sequence 4, Appl
29	204.5	18.2	429	3	PCT-US95-05523-4	Sequence 4, Appl
30	204	18.1	547	2	US-08-340-203A-3	Sequence 3, Appl
31	204	18.1	547	2	US-08-452-567-3	Sequence 3, Appl
32	204	18.1	547	2	US-08-452-567-3	Sequence 3, Appl
33	201.5	17.9	345	1	US-08-102-942A-2	Sequence 2, Appl
34	199	17.7	675	1	US-08-317-522A-9	Sequence 9, Appl
35	199	17.7	675	1	US-08-439-818A-9	Sequence 9, Appl
36	199	17.7	675	2	US-08-751-963-9	Sequence 9, Appl
37	199	17.7	675	2	US-08-738-975-9	Sequence 9, Appl
38	199	17.7	675	2	US-08-728-626-9	Sequence 9, Appl
39	194	17.2	449	1	US-08-102-942A-6	Sequence 6, Appl

KW Alzheimer's disease; therapy; diagnosis.
OS Mus sp.

FT Key Location/Qualifiers
FH Region 119..139
FT /note= "zinc finger motif"
FT Region 147..167
FT /note= "zinc finger motif"
FT Region 175..195
FT /note= "zinc finger motif"
FT Region 203..224
FT /note= "zinc finger motif"
FT Region 460..480
FT /note= "zinc finger motif"
FT Region 491..513
FT /note= "zinc finger motif"

CA2194256-A.

05-MAR-1998.

02-JAN-1997; 194256.

PR 05-SEP-1996; US-711417.

PA (GEOH) GEN HOSPITAL CORP.

PI Georgopoulos K;

DR WPI: 98-378292/33.

DR N-PSDB: V42808.

PT New nucleic acid encoding Ikaros protein involved in early differentiation of lymphocytes - existing in several isoforms, and related products, used to treat e.g. immune diseases or cancer and to control cell differentiation

PS Claim 7: Page 75-77; 158pp; English.

CC This is the amino acid sequence of murine Ikaros protein isoform mtk-1, deduced from a cDNA clone (see V42808) obtained from a mature murine T cell line E14 library. Native Ikaros is active in the early stages of lymphocyte differentiation, binding to and activating the CD3-delta gene enhancer (see V42804). Proteins of the murine Ikaros family (see also W70963 and W70965-68) are isoforms that arise from differential splicing of Ikaros gene transcripts, and contain different combinations of zinc fingers.

CC They are expressed primarily in T cells in the adult and may play a role as a genetic switch regulating entry into the T cell lineage. The murine and human sequences (see W70964, W70969 and W70971) are very similar. The invention provides Ikaros nucleic acids, vectors and host cells expressing Ikaros polypeptides. These can be used to treat T and B cell diseases (e.g. immune deficiencies caused by drugs, radiation or cancer), to control expression of heterologous genes placed under control of an Ikaros-responsive element, to treat nervous system diseases (e.g. Alzheimer's disease) and to modulate cell division, amplification or differentiation, especially in haematopoietic cells. Some Ikaros isoforms are antagonistic of others and may be used to inhibit interaction with DNA sequences.

CC Sequence 518 AA;

Query Match 59.4%; Score 669; DB 1; Length 518;

Best Local Similarity 73.7%; Pred. No. 4.6e-59;

Matches 126; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

Qy 47 RPTSGKMNCDVGLSCISFNVLWVHKRSHTGERPFQCNQCGASFTQGNLLRHKLHTGE 106

Db 111 RLPNGKLCDCIGVICGPNVLWVHKRSHTGERPFQCNQCGASFTQGNLLRHKLHSGE 170

Qy 107 KPFKCHLCNYACORRDALTGLHRLTHSVKPYKCEFCGRSYKQRSLEHKKERCRTFLOST 166

Db 171 KPFKCHLCNYACRRDALTGLHRLTHSVKPYKCEFCGRSYKQRSLEHKKERCRTFLOST 230

Qy 167 D-PG-----DTASAEARHIKAEMGSRALVLDRLASNVAKRKSSMPQKF 209

Db 231 GLPGVCPVKEETHNHNMAEDLCKIGAERSLVLDRLASNVAKRKSSMPQKF 281

RESULT 15

R92014

ID R92014 standard; Protein; 431 AA.

AC R92014; 1996

DT 08-MAY-1996 (first entry)

DE Murine Ikaros protein mtk-2.
KW Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte; immunocomprised; immune system disorder; nervous system disorder;
KW animal model; mtk-2.
OS Mus musculus.
PN WO9604372-A1.
PD 15-FEB-1996.
PR 28-JUL-1994; US-283300.
PR (GEOH) GEN HOSPITAL CORP.
PA Georgopoulos K;
PI Georgopoulos K;
DR WPI: 96-129389/13.
DR N-PSDB: T16059.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine effects of treatment for immune and nervous system disorders
PT Disclosure; Fig 1; 102pp; English.
CC The sequence of 48 kba mouse Ikaros protein mtk-2 (R92014) was deduced from mouse Ikaros cDNA (T16059) isolated from a mature T-cell line E15 library. Ikaros protein is a master regulator of hematopoietic differentiation and a major determinant in lymphocyte differentiation. Other isoforms of Ikaros (see R92016-19) arise from differential splicing of Ikaros gene transcripts. Transgenic animals, esp. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.

CC Sequence 431 AA;

Query Match 49.2%; Score 554; DB 1; Length 431;

Best Local Similarity 75.2%; Pred. No. 1.2e-47;

Matches 106; Conservative 11; Mismatches 16; Indels 8; Gaps 2;

Qy 77 GERPFQCNQCGASFTQGNLLRHKLHTGEKPKCHLCNYACORRDALTGLHRLTHSVKPK 136

Db 54 GERPFQCNQCGASFTQGNLLRHKLHTGEKPKCHLCNYACRRRDALTGLHRLTHSVKPK 113

Qy 137 YKCEFCGRSYKQRSLEHKKERCRTFLOSTD-PG-----DTASAEARHIKAEMGSE 188

Db 114 HKGCGCRSYKQRSLEHKKERCRTFLOSTD-PG-----DTASAEARHIKAEMGSE 173

Qy 189 LVLDRLASNVAKRKSSMPQKF 209

Db 174 LVLDRLASNVAKRKSSMPQKF 194

Search completed: November 6, 1999, 06:55:01

Job time: 83 sec

RESULT 12
 ID R2017 standard; Protein; 518 AA.
 AC R2017;
 DT 09-MAY-1996 (first entry)
 DE Marine Ikaros protein m1k-1.
 KW Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
 KW immunocompromised; immune system disorder; nervous system disorder;
 KW animal model; m1k-1.
 OS Mus musculus.
 PH Key Location/Qualifiers
 FT domain 119..140
 FT /label= F1
 FT /note= "zinc finger domain F1"
 FT domain 147..167
 FT /label= F2
 FT /note= "zinc finger domain F2"
 FT domain 175..195
 FT /label= F3
 FT /note= "zinc finger domain F3"
 FT domain 203..224
 FT /label= F4
 FT /note= "zinc finger domain F4"
 FT domain 460..480
 FT /label= F5
 FT /note= "zinc finger domain F5"
 FT domain 491..513
 FT /label= F6
 FT /note= "zinc finger domain F6"
 DT 09-04-1972-AL.
 PD 15-FEB-1996.
 PF 28-JUL-1995; U09345.
 PR 29-JUL-1994; US-283300.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 96-129389/13.
 DR N-PSDB: T016062.
 PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine PT effects of treatment for immune and nervous system disorders
 PS Disclosure: Fig 4; 102pp; English.
 CC The sequence of 57.5 kDa mouse Ikaros protein m1k-1 (R92017) was deduced from mouse Ikaros cDNA (T16062) isolated from a mature T-cell line E15 library. Ikaros protein is a master regulator of hematopoietic differentiation and a major determinant in lymphocyte differentiation. Other isoforms of Ikaros (see R92014, R92016 and R92018-19) arise from differential splicing of Ikaros gene transcripts. Transgenic animals, esp. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of CC treatments for immune or nervous system disorders.
 CC Sequence 518 AA;
 SQ

Query Match 59.4%; Score 669; DB 1; Length 518;
 Best Local Similarity 73.7%; Pred. No. 4.6e-59;
 Matches 126; Conservative 15; Mismatches 22; Indels 8; Gaps 2;
 QY 47 RPTSGKMNCDVGLSCISFNVLVHVKRSHGTPFCNOCGASFTQGNLRLHIKLTGE 106
 Db 111 RLPNGKLKDCIGVICGPNVLMVHVKRSHGTPFCNOCGASFTQGNLRLHIKLTGE 170
 QY 107 KPFKCHLCNACRRDALTGHLRTHSVKPKCEFCGRSYKQSSLEEHKRCRTFLQST 166
 Db 171 KPFKCHLCNACRRDALTGHLRTHSVKPKCEFCGRSYKQSSLEEHKRCRTFLQST 230
 QY 167 D-PG-----DTASAEARHKAEMGSEALVLDRLASNVAKRKSSMPQKF 209
 Db 231 GLPGVCPVKEETNHNEMAEIDLCKIGAERSLVLDRLASNVAKRKSSMPQKF 281

RESULT 13
 ID W70966 standard; Protein; 518 AA.
 AC W70966;
 DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform m1k-1.
 KW Ikaros; m1k-1; transcription factor; mouse; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;

W72674
 ID W72674 standard; Protein; 518 AA.
 AC W72674;
 DT 14-JAN-1999 (first entry)
 DE Mouse Ikaros m1k-1.
 KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
 KW differentiation marker; immune system; corpus striatum; AIDS;
 OS Mus sp.
 PN US5824770-A.
 PD 20-OCT-1998.
 PF 05-JUN-1995; 465590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 98-582621/49.
 DR N-PSDB: V66971.
 PT Ikaros poly:peptide(s) - useful for treating disorders of immune system or corpus striatum
 PS Claim 1: Column 61-66; 111pp; English.
 CC The present invention describes a purified peptide having at least one of the following properties: (a) it stimulates transcription of a DNA sequence under the control of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (c) it competitively inhibits the binding of a naturally occurring Ikaros isoform to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (d) it competitively inhibits Ikaros binding to Ikaros responsive elements; or (e) it inhibits protein-protein interactions of transcriptional complexes formed with naturally occurring Ikaros isoforms. The proteins, provided that they stimulate gene transcription under the control of delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit binding of naturally occurring Ikaros isoforms to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit Ikaros binding to Ikaros-responsive elements and/or inhibit protein-protein interactions of transcriptional complexes with naturally occurring Ikaros isoforms, can be used to treat immune system disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g. Alzheimer's disease. The present sequence represents a specifically claimed mouse Ikaros protein.
 CC Sequence 518 AA;
 SQ

Query Match 59.4%; Score 669; DB 1; Length 518;
 Best Local Similarity 73.7%; Pred. No. 4.6e-59;
 Matches 126; Conservative 15; Mismatches 22; Indels 8; Gaps 2;
 QY 47 RPTSGKMNCDVGLSCISFNVLVHVKRSHGTPFCNOCGASFTQGNLRLHIKLTGE 106
 Db 111 RLPNGKLKDCIGVICGPNVLMVHVKRSHGTPFCNOCGASFTQGNLRLHIKLTGE 170
 QY 107 KPFKCHLCNACRRDALTGHLRTHSVKPKCEFCGRSYKQSSLEEHKRCRTFLQST 166
 Db 171 KPFKCHLCNACRRDALTGHLRTHSVKPKCEFCGRSYKQSSLEEHKRCRTFLQST 230
 QY 167 D-PG-----DTASAEARHKAEMGSEALVLDRLASNVAKRKSSMPQKF 209
 Db 231 GLPGVCPVKEETNHNEMAEIDLCKIGAERSLVLDRLASNVAKRKSSMPQKF 281

RESULT 14
 ID W70966 standard; Protein; 518 AA.
 AC W70966;
 DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform m1k-1.
 KW Ikaros; m1k-1; transcription factor; mouse; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;

[illegible]


```

Db 116 KPFKCHLCNACRRDALTGHLRTHSVGKPHKCGYGRSYKQRTSLEEHKERCNYLES 175
QY 167 D-PG-----DTASAEARHIKAEMSERALVLDRLASNVAKRKSMPQKF 209
DB 176 GLPGTLPVKEETHKSEMAEDLCKIGSERSLVLDRLASNVAKRKSMPQKF 227

RESULT 5
W72672
LU W/2672 standard; Protein; 461 AA.
AC W72672; 1999 (first entry)
DE Human Ikaros.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; Immune system; corpus striatum; AIDS;
KW Alzheimer's disease.
OS Homo sapiens.
PN US5824770-A.
PD 20-OCT-1998.
PF 05-JUN-1995; 465590.
PR 02-MAY-1994; US-238212.
PR 14-SEP-1992; US-946233.
PR 14-SEP-1993; US-121438.
PR 05-JUN-1995; US-465590.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-582621/49.
DR N-PSDB: V66969.
PT Ikaros poly.peptide(s) - useful for treating disorders of immune
PT system or corpus striatum
PS Claim 1; Column 55-58; ililpp; English.
CC The present invention describes a purified peptide having at least one
CC of the following properties: (a) it stimulates transcription of a DNA
CC sequence under the control of a delta A element, an NFkB element or an
CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
CC consensus sequence; (c) it competitively inhibits the binding of a
CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC (e) it inhibits protein-protein interactions of transcriptional complexes
CC formed with naturally occurring Ikaros isoforms. The proteins, provided
CC that they stimulate gene transcription under the control of delta A
CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit binding of naturally occurring Ikaros isoforms to
CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC inhibit protein-protein interactions of transcriptional complexes with
CC naturally occurring Ikaros isoforms, can be used to treat immune system
CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC Alzheimer's disease. The present sequence represents a specifically
CC claimed human Ikaros protein.
SQ Sequence 461 AA;

Query Match 60.6%; Score 682.5; DB 1; Length 461;
Best Local Similarity 73.8%; Pred. No. 1.8e-60;
Matches 127; Conservative 16; Mismatches 20; Indels 9; Gaps 2;

QY 47 RPTSGKMNDCVGLSCISFNVLVWVHKRSHGTGERPFCNQCAGSTQKGNLLRHLKHTGE 106
DB 56 RLPNGKLKDCIGICIGPNVLVWVHKRSHGTGERPFCNQCAGSTQKGNLLRHLKHTGE 115
QY 107 KPFKCHLCNACRRDALTGHLRTHSVGKPHKCGYGRSYKQRTSLEEHKERCRTFLOST 166
DB 116 KPFKCHLCNACRRDALTGHLRTHSVGKPHKCGYGRSYKQRTSLEEHKERCNYLES 175
QY 167 D-PG-----DTASAEARHIKAEMSERALVLDRLASNVAKRKSMPQKF 209
DB 176 GLPGTLPVKEETHKSEMAEDLCKIGSERSLVLDRLASNVAKRKSMPQKF 227

```

```

RESULT 6
W70964
ID W70964 standard; Protein; 461 AA.
AC W70964;
DE Human Ikaros isoform hik-1.
KW Ikaros; hik-1; transcription factor; human; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
KW Alzheimer's disease; therapy; diagnosis.
OS Homo sapiens.
PN Location/Qualifiers
FH Key 64..84
FT Region /note= "zinc finger motif"
FT Region 92..112 /note= "zinc finger motif"
FT Region 120..140 /note= "zinc finger motif"
FT Region 148..169 /note= "zinc finger motif"
FT Region 406..426 /note= "zinc finger motif"
FT Region 434..456 /note= "zinc finger motif"
FT CA2194256-A.
PN 05-MAR-1998.
PD 02-JAN-1997; 194256.
PR 05-SEP-1996; US-711417.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR N-PSDB: V42806.
PT New nucleic acid encoding Ikaros protein involved in early
PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer and
PT to control cell differentiation
PS Claim 7; Page 70-72; 159pp; English.
CC This is the amino acid sequence of human Ikaros protein isoform
CC hik-1, deduced from a cDNA clone (see V42806) obtained from a
CC Jurkat T cell line cDNA library. Native Ikaros is active in the
CC early stages of lymphocyte differentiation, binding to and
CC activating the CD3-delta gene enhancer (see V42804). Proteins
CC of the human Ikaros family (see also W70969 and W70971) are
CC isoforms that arise from differential splicing of Ikaros gene
CC transcripts, and contain different combinations of zinc fingers.
CC They are expressed primarily in T cells in the adult and may play a
CC role as a genetic switch regulating entry into the T cell lineage.
CC The human and murine sequences (see also W70963 and W70965-68) are
CC very similar. The invention provides Ikaros nucleic acids, vectors
CC and host cells expressing Ikaros polypeptides. These can be used
CC to treat T and B cell diseases (e.g. immune deficiencies caused by
CC drugs, radiation or cancer), to control expression of heterologous
CC genes placed under control of an Ikaros-responsive element, to
CC treat nervous system diseases (e.g. Alzheimer's disease) and to
CC modulate cell division, amplification or differentiation, especially
CC in haematopoietic cells. Some Ikaros isoforms are antagonistic of
CC others and may be used to inhibit interaction with DNA sequences.
SQ Sequence 461 AA;

```

```

Query Match 60.6%; Score 682.5; DB 1; Length 461;
Best Local Similarity 73.8%; Pred. No. 1.8e-60;
Matches 127; Conservative 16; Mismatches 20; Indels 9; Gaps 2;

QY 47 RPTSGKMNDCVGLSCISFNVLVWVHKRSHGTGERPFCNQCAGSTQKGNLLRHLKHTGE 106
DB 56 RLPNGKLKDCIGICIGPNVLVWVHKRSHGTGERPFCNQCAGSTQKGNLLRHLKHTGE 115
QY 107 KPFKCHLCNACRRDALTGHLRTHSVGKPHKCGYGRSYKQRTSLEEHKERCRTFLOST 166
DB 116 KPFKCHLCNACRRDALTGHLRTHSVGKPHKCGYGRSYKQRTSLEEHKERCNYLES 175
QY 167 D-PG-----DTASAEARHIKAEMSERALVLDRLASNVAKRKSMPQKF 209

```

```

OS Mus sp.
FH key Location/Qualifiers
FT domain 290..344
FT /label= Activation_domain
FT /note= "conserved activation domain of Aiolos"
PT WO9714714-A1.
PD 24-APR-1997.
PF 17-OCT-1996; U16774.
PR 18-OCT-1995; US-005529.
PA 14-MAY-1996; US-017646.
PA (GEO ) GEN HOSPITAL CORP.
PI Georgopoulos K, Morgan BA;
PI WPI: 97-245047/22.
DR N-PSDB; T60490.
PT Aiolos polypeptide and corresponding DNA - used to reconstitute a
PT mammalian immune system, for the treatment of T cell leukaemia(s)
PT and lymphoma(s)
PS Disclosure: Page 76-78: 115pp; English.
CC Mouse Aiolos (W1574) is a homologue of Ikaros whose expression is
CC restricted to lymphoid lineage. Its amino acid sequence was
CC deduced from a cDNA clone (T60490) obtd. from a mouse spleen cDNA
CC library. A human Aiolos partial polypeptide has also been
CC identified (see also W15575). Aiolos forms homodimers that are
CC potent transcriptional activators. It also forms less potent
CC heterodimers with Ikaros. Recombinant Aiolos can be produced in
CC transformed host cells. Aiolos misexpressing immune system
CC components can be used to reconstitute the immune system of a
CC mammal. The Aiolos protein, coding sequence or cells expressing
CC Aiolos may be used to treat a disorder in an animal, esp. by gene
CC therapy. Such disorders include T-cell leukaemias and lymphomas.
CC Sequence 507 AA;
SQ

Query Match 90.6%; Score 1020.5; DB 1; Length 507;
Best Local Similarity 89.5%; Pred. No. 3e-94;
Matches 187; Conservative 12; Mismatches 9; Indels 1; Gaps 1;

QY 1 ERENVLKSPGMAEPEPTYSRYENYENIKLRHVVSDDSRPTSGKNMCDVCGL 60
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
66 DRDINIKPEPMGDAESESPYSYAREYDYESIKLERH-VPYDNRPTSGKNMCDVCGL 124
QY 61 SCTSFNVLVHKSHSGTGERPFCNCGASFTQKGNLLRHLKLTGKPKFCHLCNVACOR 120
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
125 SCTSFNVLVHKSHSGTGERPFCNCGASFTQKGNLLRHLKLTGKPKFCHLCNVACOR 184
QY 121 RDALTGLHRTSHVEKPKYKCEFCGRSYKORSLEEHEKRCRTFLQSTDPGDTASAEARHIK 180
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
185 RDALTGLHRTSHVEKPKYKCEFCGRDYKORSLEEHEKRCRAFLQNPDLGDAASVEARHIK 244
QY 181 AEMGSEALVLDRLASNVARKSSMPQKF 209
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
245 AEMGSEALVLDRLASNVARKSSMPQKF 273

RESULT 3
R46964
ID R46964 standard; Protein; 537 AA.
AC R46964;
CT 21-OCT-1994 (first entry)
PE Peptide with Ikaros protein activity.
KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
KW corpus striatum; regulatory gene.
OS Homo sapiens.
PI Georgopoulos K;
PI WPI: 96-129389/13.
DR N-PSDB; T16060.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
PT severely immuno-compromised and can be used as model to determine
PT effects of treatment for immune and nervous system disorders
PT Disclosure; Fig 2: 102pp; English.
CC An almost full-length cDNA sequence (T16060) codes for part
CC a master regulator of haematopoietic differentiation and a major
CC determinant in lymphocyte specification and development. Different
CC isoforms (see R92014 and R92016-19) of mouse Ikaros have also been
CC isolated. Transgenic animals, pref. mice, having a mutated Ikaros
CC transgene, esp. a mutation that alters the DNA binding domain of the
CC Ikaros protein, are used as models to determine the effects of
CC treatments for immune or nervous system disorders.
SQ Sequence 461 AA;

Query Match 60.6%; Score 682.5; DB 1; Length 461;
Best Local Similarity 73.8%; Pred. No. 1.8e-60;
Matches 127; Conservative 16; Mismatches 20; Indels 9; Gaps 2;

QY 47 RPTSGKMNCDVCGLSCTSFNVLVHKSHSGTGERPFCNCGASFTQKGNLLRHLKLTGGE 106
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
56 RUPNGKLKCDICIGIICIPNVLVHKSHSGTGERPFCNCGASFTQKGNLLRHLKLTGGE 115
QY 107 KPFKCHLCNVACORRDALTGLHRTSHVEKPKYKCEFCGRSYKORSLEEHEKRCRTFLQST 166
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
167 D-PG-----DTASAEARHIKAEKMGSEALVLDRLASNVARKSSMPQKF 209
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
250 GLPGTLYPVKEETKHSMAEDLKIGSERSLVLDRLASNVARKSSMPQKF 301

RESULT 4
R92015
ID R92015 standard; Protein; 461 AA.
AC R92015;
CT 09-MAY-1996 (first entry)
DE Human Ikaros protein h1k-1.
KW Ikaros; transgene; transgenic animal; transgenic mouse; h1k-1;
KW immunocomprised; immune system disorder; nervous system disorder;
KW animal model.
OS Homo sapiens.
PI Georgopoulos K;
PI WPI: 96-129389/13.
DR N-PSDB; T16060.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
PT severely immuno-compromised and can be used as model to determine
PT effects of treatment for immune and nervous system disorders
PT Disclosure; Fig 2: 102pp; English.
CC An almost full-length cDNA sequence (T16060) codes for part
CC a master regulator of haematopoietic differentiation and a major
CC determinant in lymphocyte specification and development. Different
CC isoforms (see R92014 and R92016-19) of mouse Ikaros have also been
CC isolated. Transgenic animals, pref. mice, having a mutated Ikaros
CC transgene, esp. a mutation that alters the DNA binding domain of the
CC Ikaros protein, are used as models to determine the effects of
CC treatments for immune or nervous system disorders.
SQ Sequence 461 AA;

```

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 06:55:00 ; Search time 21.23 Seconds
(without alignments)
233.179 Million cell updates/sec

Title: US-09-019-348-8
Perfect score: 126
Sequence: 1 ERDENVLKSEPMGNAEPEI.....VLDRLASNVAKRKSMPQKF 209

Scoring table: BLOSOM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1126	100.0	209	1 W15575	Human Aiolos polyp
2	1020.5	90.6	507	1 W15574	Mouse Aiolos polyp
3	682.5	60.6	537	1 R46964	Peptide with Ikaros
4	682.5	60.6	461	1 R20215	Human Ikaros prote
5	682.5	60.6	461	1 W2672	Human Ikaros. Ikar
6	682.5	60.6	461	1 W70964	Human Ikaros isofo
7	682.5	60.6	516	1 W70971	Human Ikaros isofo
8	679.5	60.3	470	1 R20201	Ikaros protein. Tr
9	679.5	60.3	470	1 W2678	Ikaros protein gen
10	679	60.3	568	1 R46965	Ikaros zinc finger
11	676.5	60.1	470	1 W70970	Ikaros isoform i c
12	669	59.4	518	1 R20217	Murine Ikaros prot
13	669	59.4	518	1 W2674	Mouse Ikaros mix-1
14	669	59.4	518	1 W70966	Mouse Ikaros isofo
15	554	49.2	431	1 R20214	Murine Ikaros prot
16	551	48.9	431	1 R46963	Murine Ikaros zinc
17	551	48.9	431	1 W2671	Murine Ikaros mIk-
18	551	48.9	431	1 W70963	Mouse Ikaros isofo
19	432	38.4	432	1 R20216	Murine Ikaros prot
20	432	38.4	432	1 W2673	Mouse Ikaros mix-3
21	432	38.4	432	1 W70965	Mouse Ikaros isofo
22	418.5	37.2	334	1 R20200	Ikaros protein. Tr
23	418.5	37.2	334	1 W2677	Ikaros isoform. Ik
24	418.5	37.2	334	1 W70969	Human Ikaros prote
25	403.5	35.8	390	1 R20218	Murine Ikaros prot
26	403.5	35.8	390	1 W2675	Mouse Ikaros mix-4
27	403.5	35.8	390	1 W70967	Mouse Ikaros isofo
28	313	27.8	56	1 W2668	Ikaros protein SEQ
29	313	27.8	56	1 W2689	Ikaros protein SEQ
30	303	26.9	56	1 W2687	Ikaros protein SEQ
31	271.5	24.1	412	1 Y70708	Renal cancer assoc
32	269	23.9	711	1 W37504	Human OTK18. Novel
33	263	23.4	532	1 W52187	Human BRCA1-associ
34	258	22.9	706	1 R68743	BCL-6 zinc finger
35	258	22.9	181	1 R83490	Zif268-Zif268 zinc
36	258	22.9	181	1 W84394	The zif268-zif268
37	256	22.7	561	1 W82404	Human SRE-ZBP anal
38	251.5	22.3	608	1 W06316	TRP-1 protein. DNA
39	249.5	22.2	803	1 W81756	Myc-binding zinc-f
40	244	21.7	421	1 W81631	Human stem cell zi
41	238.5	21.2	361	1 W81632	Human stem cell zi
42	230	20.4	88	1 W87557	A nucleic acid bin
43	230	20.4	88	1 W84299	An anti-HIV zinc i

44 230 20.4 88 1 W8390 zinc finger protel
45 227 20.2 500 1 W37948 Amino acid sequenc

ALIGNMENTS

RESULT 1
W15575
ID W15575 standard; Protein; 209 AA.
AC W15575;
DT 07-JUL-1997 (first entry)
DE Human Aiolos polypeptide (partial sequence).
KW Aiolos; transcription activator; immune system; T lymphocyte;
B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW Transgenic animal.
OS Homo sapiens.
PN W0974774-11.
PD 24-APR-1997.
PF 17-OCT-1996; U16774.
PR 18-OCT-1995; US-005529.
PR 14-MAY-1996; US-017646.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K, Morgan BA;
DR WPI; 97-245047/22.
DR N-PSDB; T60491.
PT Aiolos polypeptide and corresponding DNA - used to reconstitute a
mammalian immune system, for the treatment of T cell leukaemia(s)
PT and lymphoma(s).
PS Disclosure; Fig 5b; 115pp; English.
CC A human Aiolos partial polypeptide (W15575) is encoded by an
isolated cDNA clone (T60491). Aiolos is a homologue of Ikaros
whose expression is restricted to lymphoid lineage. It forms
CC homodimers that are potent transcriptional activators. It also
CC forms less potent heterodimers with Ikaros. Recombinant Aiolos can
CC be produced in transformed host cells. Aiolos misexpressing immune
CC system components can be used to reconstitute the immune system of
CC a mammal. The Aiolos protein, coding sequence or cells expressing
CC Aiolos may be used to treat a disorder in an animal, esp. by gene
CC therapy. Such disorders include T-cell leukaemias and lymphomas.
SQ Sequence 209 AA;

Query Match 100.0%; Score 1126; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.6e-105;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERDENVLKSEPMGNAEPEIPYSYSREYNEYENIKLERHVVSFDSSRPTSGKMNCDVCGL 60
DB 1 ERDENVLKSEPMGNAEPEIPYSYSREYNEYENIKLERHVVSFDSSRPTSGKMNCDVCGL 60

QY 61 SCIFNVLVWVKRSHHTGERPFCQCCGASTOKGNLRLHKLHTGKPEKCHLCNACOR 120
DB 61 SCIFNVLVWVKRSHHTGERPFCQCCGASTOKGNLRLHKLHTGKPEKCHLCNACOR 120

QY 121 RDALTGLRTHSVKPKCEFCGRYSKORSLEEHKERCRTFLQSTDPGDTASAEARHK 180
DB 121 RDALTGLRTHSVKPKCEFCGRYSKORSLEEHKERCRTFLQSTDPGDTASAEARHK 180

QY 181 AEMGSEALVLDRLASNVAKRKSMPQKF 209
DB 181 AEMGSEALVLDRLASNVAKRKSMPQKF 209

RESULT 2
W15574
ID W15574 standard; Protein; 507 AA.
AC W15574;
DT 07-JUL-1997 (first entry)
DE Mouse Aiolos polypeptide.
KW Aiolos; transcription activator; immune system; T lymphocyte;
B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW Transgenic animal.

THIS PAGE BLANK (USPTO)

High quality sequence stops: 313 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1126 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 313.

FEATURES

```

/organism="Homo sapiens"
/db_xref="CDB:40407"
/db_xref="taxon:9606"
/clone="IMAGE:32060"
/clone_lib="Soares infant brain lN18"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DM108 (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda BA; Site.1: Not
I; Site.2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5',
NACTGGAGAGAAATCGGGCGCGAGGAATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."
85 c 116 g 121 t
2 others

```

BASE COUNT
ORIGIN

Query Match 13.6%; Score 85.2; DB 21; Length 470;
Best Local Similarity 57.5%; Pred. No. 1.3e-15;
Matches 153; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

203	QY	taatggttcataagcgaagccatactggtgaacgcccaattccagtgtaactcagtggtggg	262
30	Db	TAATCGCGCATCAGAGAATTACACTGGGGAGAAACCTTTTGAATGTAATGAGTGTGGGA	89
263	QY	catctttactcagaaggtaaacctctccgccacattaaactgcacacaggggaaaaac	322
90	Db	GATGCTTTACTTCTAAAGAAACCTACTTGNATCATCACCGAATCCATACTCGAGAAAGC	149
323	QY	cttttaagtgtaacctctgcaactatgcatgcccagaagaagatgcgtcacggggcacc	382
150	Db	CCATCAATGTAAAGGAATGTGGGAAAGCCTTCAGTATCAATGCCAAACTAACTAGGCATC	209
383	QY	ttaggcacattctgtgagaaacccatacaatgtgatttgtgggaaggagtaccaacg	442
210	Db	AGAGGATACATATCGGGAGAAACCTTCAANTGTATGGAATGTGAGAAAGCATTCAGCT	269
443	QY	agagaagttccctcttgaggagcacaag	468
270	Db	GTAGTCTTAACATATATTGTGCACCAG	295

RESULT 12
R2287A

LOCUS	R22824	431 bp	mRNA	EST	19-APR-1995
DEFINITION	Y947a01.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:35545 5', mRNA sequence.				

ACCESSION	R22824
NID	q777649

VERSION R22824.1 GI:777649
KEYWORDS EST.

KEYWORDS	ESSI.
SOURCE	human.
ORGANISM	Homo sapiens

ORGANISM HOMO SAPIENS
Eukaryota; Metazoa;
Euthorica; Primates;

REFERENCE
1 (bases 1 to 431)
Eutheria; Primates;
Mammalia; Carnivora

AUTHORS
 Hillier, L., Clark, N.
 Holman, M., Hultman, J.

Parsons, J., Rifkin, J., Trevaskis, E., Water

Wilson, R.

TITLE	JOURNAL	COMMENT
-------	---------	---------

The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

High quality sequence stops: 318 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL : contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1116 Std Error: 0.00

seq primer: M13RPI
High quality sequence stop: 318.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

Location/Qualifiers
1. 431
/organism="Homo sapiens"
/db_xref="GDB:408046"
/db_xref="taxon:9606"
/map="1"
/clone="IMAGE:35545"
/clone.lib="Soares infant brain IN1B"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda; Site:2: Hind III; 1st strand cDNA I - oligo(dT) primer [5, AACTGTGAAGATTCGGCCGCAGGAATTTTTTTT double-stranded cDNA was ligated to H (pharmacia), digested with Not I and into the Not I and Hind III sites of Library went through one round of normal constructed by Bento Soares and M.Fat
139 a 77 c 110 q 103 t 2 ot

```

BASE COUNT
ORIGTN

Query Match 13.6%; Score 85.2; DB 22; Length 431;
Best Local Similarity 57.5%; Pred. No. 1.3e-15;
Matches 153; Conservative 0; Mismatches 113; Indels 0; Caps 0;

QY 203 taatggttcataagcgaagccatactgggtgaacgcccattccagtgtaatatcagtggtggtg 262

Db 18 TAATGGCGGCATCAGAGAAATTCACATGGGGGAGAAACCTTTTGAATGTAATGAGTGTGGGA 77

Qy 263 catctttactcagaaggtaacctctctccgcacattaaactgcacacaggggaaaaac 322

Db 78 GATGCTTTACTTCTAAAAGAAACCTACTTGTATCATCACCGAATCCATACTGGAGAAAAGC 137

Qy 323 cttttaaagtgtcacctctgcaactatgcatgccaaagaagagatgcgtcacggggcacc 382

Db 138 CCTATCAATGTANGGAATGTGGGAAGCCTTCAGTATCAATGCCAACTAAGTAGGCATC 197

QY 383 ttaggacacattctgtggaacccctacaaaatgtgagtttgtggaaggagtacaagc 442

Db 198 AGAGGTACATCTGGGGAGAAAACCTTCAAAATGTTATGGRAATGTGAGAAAGCAATTCAGCT 257

Qy 443 agagaagtcccttgaggagacaag 468

258 GTAGTTCTAACTATATTGTGACCAG 283

RESULT 13
AA119538

AA119538	LOCUS	DEFINITION
----------	-------	------------

DEFINITION

ACCESSION
NID
UNION

VERSION

AA119538 389 bp mRNA EST 17-FEB-1997
mp68a03.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:574348 5',
mRNA sequence.

AA119538
g1677047

91077047
AA119538.1 GI:1677047

Db 49 GGGCCTCATTACCCAGAGGGCAACCTGCTCGGCACATCAAGC 5

RESULT 7

R19256
LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

14-APR-1995
EST
191256
mrna
yg24h05.r1 Soares infant brain lNIB Homo sapiens cDNA clone
IMAGE:33351 5', mRNA sequence.
R19256
g772866
R19256.1 GI:772866
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 467)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfling,T., Soares,M., Tan,F., and
Trevasakis,E., Watsonson,K., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```

Email: estewartson.wusli.edu
Insert Size: 1113
High quality sequence steps: 311 Source: IMAGE Consortium, LLM
This clone is available royalty-free through LLM ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1113 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 311.
Location/Qualifiers
1. .467
/organism="Homo sapiens"
/db_xref="GDB:405698"
/db_xref="taxon:9606"
/clone="IMAGE:33351"
/clone_lib="Soares infant brain LNIb"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="PH10B (ampicillin resistant)"
/notes="Organ: whole brain; Vector: Lfam1d BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I oligo(dT) primer [5'
ACTGTGAAGATTCGGCGCCGAGGAATTTTCTTTTCTTTT 3'] ;
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfam1d BA vector.
Library went through one round of normalization. Library
constructed by Post Genomic Technology."

```

[illegible]


```
Query Match      38.2%; Score 239.6; DB 48; Length 546;
Best Local Similarity 85.3%; Pred. No. 1.7e-62;
Matches 279; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 1 gaaagagatgagaatgttttaagtcagaaccatggagaaatgcagaagagccctgaaatc 60
DB 223 GACAGAGATGAGAACATTTATGAACCGGAGCCCATGGAGATGCAGAGAGAGTGAATG 282
QY 61 ccttacagctattcaagagaataataatgaataatgaataatgaatgagagacatgtt 120
DB 283 CCTTACAGCTATGACAGAGATACAGGACTATGAACCATTAAGCTGGAGAGACA---C 339
QY 121 gtctcaatcgatagtagcaggcccaaccagtggaagatgaactcgatgtgtggatta 180
DB 340 GTGCCCTATGACAACAGCAGACATTTATGAAGCCGAGCCCATGGAGATGCAGAGAGTGAATG 282
QY 61 ccttacagctattcaagagaataataatgaataatgaataatgaatgagagacatgtt 120
DB 283 CCTTACAGCTATGACAGAGATACAGGACTATGAACCATTAAGCTGGAGAGACA---C 339
QY 121 gtctcaatcgatagtagcaggcccaaccagtggaagatgaactcgatgtgtggatta 180
DB 340 GTGCCCTATGACAACAGCAGACATTTATGAAGCCGAGCCCATGGAGATGCAGAGAGTGAATG 282
QY 181 tctgcagctattcaagagaataataatgaataatgaataatgaatgagagacatgtt 120
DB 400 TCCAGTGTATGACAGAGATACAGGACTATGAACCATTAAGCTGGAGAGACA---C 339
QY 241 tctgcagctattcaagagaataataatgaataatgaataatgaatgagagacatgtt 120
DB 460 TCCAGTGTATGACAGAGATACAGGACTATGAACCATTAAGCTGGAGAGACA---C 339
QY 301 aaactgcacacaggggaaaaaccccttt 327
DB 520 AAACGTGCACACGGGAAAAACCTTTT 546

RESULT 3
AI550402      501 bp      mRNA      EST      23-MAR-1999
LOCUS      vx13e05.x1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1264352 3'
DEFINITION      similar to TR:008900 008900 AIOLOS ; mRNA sequence.
ACCESSION      AI550402
NID      94482765
VERSION      AI550402.1 GI:4482765
KEYWORDS      EST.
SOURCE      Mus musculus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS      Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
      Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
      Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
      Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
      Waterston,R. and Wilson,R.
TITLE      The WashU-NCI Mouse EST Project 1999
JOURNAL      Unpublished (1999)
COMMENT      On Jun 22, 1998 this sequence version replaced gi:3246999.

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:666904
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
High quality sequence stop: 381.
Location/Qualifiers
1. 501
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

FEATURES
source
```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 08:16:48 ; Search time 437.25 Seconds
(without alignments)
2833.050 Million cell updates/sec

Title: US-09-019-348-7
Perfect score: 628
Sequence: 1 gaaagagagagaatgtttt.....ctcaatgcctcagaattca 628

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

EST:*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

54: em_est22:*

55: em_est23:*

56: em_est24:*

57: em_est25:*

58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	303.4	48.3	577	39	AA863970	AA863970 vx87d08.r
C 2	239.6	38.2	546	48	AI549961	AI549961 vx02e08.x
C 3	195.5	29.5	501	48	AI550402	AI550402 vx13e05.x
C 4	179.6	28.6	609	40	AA920812	AA920812 vx86f12.r
C 5	125	19.9	557	45	AI325349	AI325349 ml06e03.y
C 6	93	14.8	339	38	AA808413	AA808413 oa43h06.s
C 7	90	14.3	467	21	R19256	R19256 y924h05.r1
C 8	89.6	14.3	355	24	H83408	H83408 y924h05.r1
C 9	89.4	14.2	509	27	AA027561	AA027561 ml06e03.r
C 10	85.6	13.6	291	32	AA334780	AA334780 EST39041
C 11	85.2	13.6	470	21	R17304	R17304 y912d09.r1
C 12	85.2	13.6	431	22	R22824	R22824 y947a01.r1
C 13	84.6	13.5	389	28	AA119538	AA119538 mp68a03.r
C 14	84.4	13.4	587	48	AI559335	AI559335 tg43h01.x
C 15	83.8	13.3	382	33	AA400326	AA400326 zu69b10.r
C 16	83.6	13.3	525	29	AA130717	AA130717 zol3h06.s
C 17	83.4	13.3	644	22	H05551	H05551 y175b08.r1
C 18	83	13.2	805	49	AI648952	AI648952 uk33c08.x
C 19	82.8	13.2	332	29	AA171117	AA171117 ms55b10.r
C 20	82.8	13.2	512	46	AI406535	AI406535 EST234821
C 21	82.8	13.2	406	54	HSW009139	AI044289 Homo sapi
C 22	82.6	13.2	837	53	HSW007817	AI042967 Homo sapi
C 23	82	13.1	494	39	AA848430	AA848430 EST191190
C 24	81.8	13.0	704	53	HSW007214	AI042364 Homo sapi
C 25	80.8	12.9	645	36	AA625515	AA625515 af72e07.r
C 26	80.8	12.9	544	44	AI243713	AI243713 qb88f01.x
C 27	80.2	12.8	758	36	AA628972	AA628972 af28h12.s
C 28	80.2	12.8	791	38	AA772346	AA772346 ai43c08.s
C 29	80.2	12.8	768	38	AA789218	AA789218 aj27c09.s
C 30	80.2	12.8	723	40	AA972741	AA972741 op90h05.s
C 31	79.8	12.7	424	26	W19169	W19169 zb25b08.r1
C 32	79.8	12.7	369	39	AA814418	AA814418 nz12d08.s
C 33	79.6	12.7	601	34	AA455712	AA455712 aa22e04.r
C 34	79.2	12.6	774	36	AA628881	AA628881 af30a12.s
C 35	78.8	12.5	662	26	W28371	W28371 46a4 Human
C 36	78.8	12.5	491	35	AA548358	AA548358 nk41a05.s
C 37	78.8	12.5	545	36	AA624117	AA624117 vg75g07.s
C 38	78.6	12.5	447	23	R98364	R98364 HH210SF Hum
C 39	78.6	12.5	447	23	R98368	R98368 HH211SF Hum
C 40	78.6	12.5	761	40	AA917951	AA917951 ol68g11.s
C 41	78.4	12.5	685	26	W27601	W27601 35a3 Human
C 42	78.4	12.5	474	28	AA113952	AA113952 zm81c11.r
C 43	78.4	12.5	465	42	AI128612	AI128612 ga62a10.s
C 44	78.4	12.5	492	44	AI311718	AI311718 g092f08.x
C 45	78.2	12.5	378	21	T63663	T63663 yc23a03.r1

ALIGNMENTS

RESULT 1
AA863970/c
LOCUS vx87d08.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1282191 5'
DEFINITION similar to TR:008900 O08900 AIOLOS ;, mRNA sequence.
ACCESSION AA863970
NID 92956449
VERSION AA863970.1 GI:2956449

AA863970 577 bp mRNA EST 11-MAR-1998
vx87d08.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1282191 5'
similar to TR:008900 O08900 AIOLOS ;, mRNA sequence.

Query Match 11.5%; Score 72.8; DB 4; Length 2031;
Best Local Similarity 53.5%; Pred. No. 2.6e-14;
Matches 152; Conservative 0; Mismatches 152; Indels 0

0.

1735	GGAGAGAGCCCTCATTATGATGCTATACCCATCTCCACAGAGACCTTGAAGAGTACATCTGGGGCT	1819
DD		
QY	403 aaacccatacaatgtgagttttgtgggaaggaggttacaaagcagag	446
DB	1819 AAGCCCTTACAAGTGTTCAGAGACTGTGGAAGAAGCCCTTCGCCACAG	1862

Search completed: November 6, 1999, 06:57:40
Job time: 141 sec

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09345
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,300
FILING DATE: 29-JULY-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MGP-027PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1296
PCT-US95-09345-3

Query Match 23.7%; Score 148.6; DB 5; Length 1296;
Best Local Similarity 73.4%; Pred. No. 3e-39; Indels 0; Gaps 0;
Matches 190; Conservative 0; Mismatches 69;
Qy 146 ccagtggaagatgaactgcgatgtgtggtattatctctgcatcagcttcaatgtcttaa 205
Db 338 CTACGGAACAACTAAAGTGTATATCTGTGGATCGTTTGATCGGGCCCAATGTGCTCA 397
Qy 206 tgggtcataagcgaagccatactggtgaacgccccattccagtgttaatacagtggtgggcat 265
Db 398 TGGTTTCAAAAAGAGTATATCTGTGTGAACGGCCTTTCCAGTGAACCAAGTCTGGGGCT 457
Qy 266 cttttactcaagaagtgtaacctctctcccccattataaactcacacaggggaaacatt 325
Db 458 CTTTACCAGAAAGGAACTCTCTCGGCACATCAAGCTGCACATCGGGTGAAGAGCCCT 517
Qy 326 ttaagtgtaacctctgcaactatgcatgccaaagaagagatgcgttcacggggcatctta 385
Db 518 TCAATGCACTCTTTGCAACTATGCTGCGCCGAGGAGGCCCTCACCGGCCACCTGA 577
Qy 386 ggacacattctgtggagaa 404
Db 578 GGACGCACTCGGAGACAA 596
RESULT 11
US-08-465-590-6
Sequence 6, Application US/08465590
Patent No. 5824770
GENERAL INFORMATION:
APPLICANT: Georgopoulos, Katia A.
TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE

NUMBER OF SEQUENCES: 164
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MGP-006C2DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1170 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1170
US-08-465-590-6
Query Match 17.1%; Score 107.6; DB 3; Length 1170;
Best Local Similarity 77.1%; Pred. No. 7.6e-26;
Matches 131; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
Qy 229 ggtgaagccattccagtgtaatacagtggtgggcatcttttactcagaaggaacctc 288
Db 160 GGTCAACGGGCTTTCCAGTGAACCAAGTCTGGGGCCTCTTACCCAGAAAGGCAACCTC 219
Qy 289 ctccgcacacattaaactgcacacaggggaaacacacacacacacacacacacacacacac 348
Db 220 CTGGGACACATCAAGTGCACATCGGGTGAAGGCCCTTCAAAATGCCATCTTTGCAACTAT 279
Qy 349 gcatgcaaaagaagatgcgttcacggggcatctttagtgacacacacacacacacacacac 398
Db 280 GCCTGCCGCGGAGGAGCGCCCTCACCGGCCACCTGAGGAGCGCACTCCGT 329
RESULT 12
PCT-US95-09345-5
Sequence 5, Application PC/TUS9509345
GENERAL INFORMATION:
APPLICANT: Georgopoulos, Katia A.
TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON

NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MGP-027PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1788 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 223..1515
PCT-US95-09345-1

Query Match 27.1%; Score 170; DB 5; Length 1788;
Best Local Similarity 67.1%; Pred. No. 3.5e-46;
Matches 283; Conservative 0; Mismatches 115; Indels 24; Gaps 2;
QY 229 ggtgaagccattccagtgtaatacagtggtgggcatcttttactcagaaggttaacctc 288
DB 382 GGTGAAGCGCTTCCAGTGCACAGTCTGGGGCCCTCTTACCCAGAAAGGCAACCTC 441
QY 289 ctccgcacattaaactgcacacaggggaaacaccttttaagtgtcacctctgcaactat 348
DB 442 CTGCGGCACATCAAGTGCACCTCGGTGGAGAGCCCTTCAATGCCATCTTTGCAACTAT 501
QY 349 gcatgccaaagaagatgcgctcacggggcatctttaggacacattctgtgggaaaccc 408
DB 502 GCCTGCCGCGGGAGGAGCGCCCTCACCGGCACCTGAGGAGCGCACTCCGTTGGTAAGCCT 561
QY 409 tacaattgtagtttttgaagagagtataacagcagaagttcccttgaggagacaaag 468
DB 562 CACAAATGTGATATTGTGGCGGAGGTATAACAGCGGAAGCTTTTAGAGGAGCATAAA 621
QY 469 gagcgctgcggtacattcttcagagcactgacc-----caggggacactgcaagtgcg 522
DB 622 GAGCGATGCCACACTACTTGGAAAGCATGGCCCTTCGGGCGGTGTGCCAGTCAATTAAG 681
QY 523 gaggcagacacataaag-----cagagatgggaagtgaagagct 564
DB 682 GAAGAAACTRACCAACACAGAGATGGCAGAGACCTGTGCAAGATAGGAGCAGAGAGGTCC 741
QY 565 ctctactgacagatagcaagcaatgtggcaaaacgaaacacacacacacacacacacac 624
DB 742 CTGTCTGGACAGGTGGCAAGCAATGTCCGCAACGTAAGAGCTCTATGCCTCAGAAA 801
QY 625 tt 626
DB 802 TT 803

RESULT 9
US-08-465-590-4
Sequence 4, Application US/08465590
Patent No. 5824770
GENERAL INFORMATION:
APPLICANT: Georgopoulos, Katia A.
TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
NUMBER OF SEQUENCES: 164
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM: Floppy disk
MEDIUM TYPE:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MGP-006C2DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1296
US-08-465-590-4

Query Match 23.7%; Score 148.6; DB 3; Length 1296;
Best Local Similarity 73.4%; Pred. No. 3e-39;
Matches 190; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 146 ccagtggaagatgaactgcgagtgtgtggattatctctgcacatcagcttcaatgtcttaa 205
DB 338 CTAACGGAAACTAAAGTGTGATATCTGGGATCTGTTGCATCGGGCCCAATGTGCTCA 397
QY 206 tggttcataagcagaccatactactgttgaaagccacattccagtgtaatacagtggtgggcat 265
DB 398 TGGTTCACAAAGAAAGTCATACTGGTGAACGGCCTTTCCAGTGCACCACTCTGTGGGCT 457
QY 266 ctcttactcagaaggttaacctctccgcacacattaaactgacacagaggggaaaaacctt 325
DB 458 CCTTACCCAGAAAGCAACCTCTCGCGCACATCAAGCTGCACCTCGGTGTGAGAAGCCCT 517
QY 326 ttaagtgtcacctctgcaactatgcatgccaagaagagagatgcgtcacaggggcatctta 385
DB 518 TCAATGCCATCTTGCACACTATGCTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 577
QY 386 ggacacattctgtggagaa 404
DB 578 GGACGCACTCCGGAGACAA 596

RESULT 10
PCT-US95-09345-3
Sequence 3, Application PC/TUS9509345
GENERAL INFORMATION:
APPLICANT: Georgopoulos, Katia A.
TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA

; LOCATION: 223..1515
US-08-465-590-2

Query Match 27.1%; Score 170; DB 3; Length 1788;
Best Local Similarity 67.1%; Pred. No. 3.5e-46;
Matches 283; Conservative 0; Mismatches 115; Indels 24; Gaps 2;

Qy	229	ggtgaagcccccattccagtgtaatacagtgtaggggacattcttttactcagaaaggttaacctc	288
Db	382	GGTGAACGGGCTTTCCAGTGCACACCAAGTCTGGGGCTCTCTTTACCCAGAAAGGCAACCTC	441
Qy	289	ctcgcacataaactgcacacaggggaaaaaccttttaagtgtcacctctgcacactat	348
Db	442	CTCGGCACATCAAGCTGCACTCGGTGAGAACCCCTTCNAATGCCATCTTTGCACTAT	501
Qy	349	gcattccaaagaagatgcgtctacaggggcatcttaggacacattctgtggaaaaacc	408
Db	502	GCCTGCCCGGAGGACGCCCTCACCGGCCACTGAGGACGCATCCGTTGTAAGCCT	561
Qy	409	tacaaatgtagtttgtgaagaggttaacacagacagaagttcccttaggagacacaag	468
Db	562	CACAAATGTGGATATTGTGCCGGAGCTATAACACGGAAGTCTTTAGAGAGGATATA	621
Qy	469	gagcgtcgctacattcttcagacactgacc-----caggggacactgcagtgcc	522
Db	622	GAGCGATGCCACAACACTACTTTGGAAGCATGGGCTTCCGGCGGTGTCACGTCATTAG	681
Qy	523	gaggcaagacacataaag-----cagagatgggaagtgaagagct	564
Db	682	GAAGAAACTAACCAACACAGATGGCAAGAGACTCTGTCAAGATAGGACGACAGAGTCC	741
Qy	565	ctcgtactgcacagattagcaagaatgtggcaaaacgaaaaagctcaatgctcagaaa	624
Db	742	CTTGTCTCTGACAGGCTGGCAACCAATGTCCGCCAAACGTTAGAGCTCTATGCCCTCAGAAA	801
Qy	625	tt	626
Db	802	TT	803

```

RESULT          7
PCT-US93-08743-2
; Sequence 2, Application PC/TUS9308743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08743
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,233
; FILING DATE: 14-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1788 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 223..1515
PCT-US93-08743-2

```

Query Match	27.1%;	Score 170;	DB 5;	Length 1788;
Best Local Similarity	67.1%;	Pred. No. 3.5e-46;		
Matches 283; Conservative	0;	Mismatches 115;	Indels 24;	Gaps 2;

Qy	229	ggtgaacgcccattccagtgtaatacgtgtgaggacatctttttactcagaaagtgaaacctc	298
Db	382	GGTGAACGGGCTTTCCAGTGCACCACTGCTGGGGCTCTTTTACCAGAAAGGCAACCTC	441
Qy	289	ctcgcgcacattaaactgcacacaggggaaaaacctttaagtgtcacctctcgaactat	348
Db	442	CTGGCGGCACATCAAGCTGCACTCGGTGAGAACCCCTTCAATGCCATCTTTGCCAACTAT	501
Qy	349	gcattgcaaaagagagatgcgctcacgggggcatcttaggacacattctgtggagaaaccc	408
Db	502	GCTTGCCTCGGAGGAGCGCCCTCACCGGCCACTCTGAGGACGCACCTCCGTTGTAAGCCT	561
Qy	409	tacaaatgtgagttttgtgaaaggagttacaacacagagaagttcccttgaggagacaag	468
Db	562	CACAAATGTGGATATTGTGCCGGAGCTATAACACGCGAAGCTCTTTAGAGGAGCATAA	621
Qy	469	gagcgctgcggtacattcttcagagcactgacc-----caggggacacactgcaagtgcg	522
Db	622	GAGCGATGCCACAACTACTTGGAAAGCATGGCGCTTCCGGCGGTGTGCCACGTCAATTAG	681
Qy	523	gaggcaagacacatacaag-----cagagatgggaagtgsaaagact	584
Db	682	GAAGAAACTAACCACACAGAGATGGCAGAGACTGTGTCAAGATAGGAGCAGAGAGTCC	741
Qy	565	ctctactgcacagattagcaaacatgtgcaaaacgaaaaagctcaatgcttcagaa	624
Db	742	CTTCTCTCGACAGGCTGGCAGCAATGTCCGCCAAACGTTAAGAGCTCTATGCCCTCAGAA	801
Qy	625	tt 626	
Db	802	tt 803	

RESULT 8
PCT-US95-09345-1
Sequence 1, Application PC/TUS9509345
GENERAL INFORMATION:
APPLICANT: Georgopoulos, Katia A.
TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09345
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,300
FILING DATE: 29-JULY-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:

; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1386 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1386
 ; PCT-US95-09345-2

Query Match 34.5%; Score 216.4; DB 5; Length 1386;
 Best Local Similarity 67.9%; Pred. No. 1.9e-61;
 Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

QY 146 ccagtggaagatgaactgcgagtgtgtgtgattatctctgcatcagcttcaatgtcttaa 205
 DB 173 CTACGGAAACTAAAGTGTGATATCTGTGGGATCATTTGGATCGGGCCCAATGTGCTCA 232
 QY 206 tgggttcataagcgaagccatactgtgaaccccatccagtgtaatacagtggtggcat 265
 DB 233 TGGTTTCAAAAGAACCCACATCGGAGAACGGCCCTTCCAGTGCATCAGTGGGGGCT 292
 QY 266 cttttactcagaagaagttaacctctccgccacattaaactgcacacaggggaaaaaacctt 325
 DB 293 CATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAGCCCT 352
 QY 326 ttaagtgtaacctctgcactatgatcccaaaagaagatgcgctcacggtggcatctta 385
 DB 353 TCAAAATGCCACCTCTGCAACTACGGCTCGCGCGGAGGAGCCCTCACTGGCCACCTGA 412
 QY 386 ggacacattctgtgagaacctcaaatgtgattgtgattgtggaaggattacaagcaga 445
 DB 413 GGACGCACTCCGTTGGTAAACCTCAAAATGTGGATATTGTGGCGAAGCTATAAACAGC 472
 QY 446 gaagttcccttgagagcagaagggcgctgcgtacattcttcagagcactgacc--c 503
 DB 473 GAACGCTTTAGAGGAACAATAAGAGCGCTGCCACAACACTATTGGAAAGCATGGCCCTTC 532
 QY 504 aggggacactgcagtgcgaggcgaagacacatcaaaagca----- 544
 DB 533 CGGGCACACTGTACCCAGTCAATTAAGAAAGAACTAAGCACAGTGAATGGCAGAAGACC 592
 QY 544 ----gagatgggaagtgaagagctctctgtactggacagattagcaagaatgtggcaa 598
 DB 593 TGTGCAAGATAGGATCAGAGAGATCTCTGTGCTGGACAGAGTAGCAAGTAATGTGCGCA 652
 QY 599 aacgaaaaagctcaatgctcagaatt 626
 DB 653 AAGCTAAGAGCTCTATGCTCAGAAATT 680

RESULT 4
 US-08-465-590-5
 ; Sequence 5, Application US/08465590
 ; Patent No. 5824770
 ; GENERAL INFORMATION:
 ; APPLICANT: Georgopoulos, Katia A.
 ; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
 ; NUMBER OF SEQUENCES: 164
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHVE & COCKFIELD
 ; STREET: 60 STATE STREET, Suite 510
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM: Floppy disk
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII (text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/465,590
 ; FILING DATE: 05-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/238,212
 ; FILING DATE: 02-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/121,438
 ; FILING DATE: 14-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/946,233
 ; FILING DATE: 14-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Paul L.
 ; REGISTRATION NUMBER: 35,695
 ; REFERENCE/DOCKET NUMBER: MPG-006C2DV
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2049 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 223..1776
 ; US-08-465-590-5

Query Match 33.4%; Score 209.8; DB 3; Length 2049;
 Best Local Similarity 67.1%; Pred. No. 3.4e-59;
 Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;

QY 146 ccagtggaagatgaactgcgagtgtgtgtgattatctctgcatcagcttcaatgtcttaa 205
 DB 560 CTACGGAAACTAAAGTGTGATATCTGTGGGATCGTTTCCAGTGCACACAGTCTGGGCTCA 619
 QY 206 tgggttcataagcgaagccatactactgtggaaccccatccagtgtaatacagtggtggcat 265
 DB 620 TGGTTCAAAAGAAAGTGCATATGTTGTCAGTGCACGGCCCTTTCAGTGCACACAGTCTGGGCT 679
 QY 266 cttttactcagaagaagttaacctctccgccacattaaactgcacacaggggaaaaaacctt 325
 DB 680 CTTTACCCAGAAAGGCAACCTCCTGCGGCACATCAAGCTGCACCTCGGGGTGAGAAGCCCT 739
 QY 326 ttaagtgtaacctctgcaactatgatcccaaaagaagatgcgctcacggtggcatctta 385
 DB 740 TCAATGCCATCTTTGCAACTATGCTTCGCGGAGGAGCGCCCTCACGGCCACCTGA 799
 QY 386 ggacacattctgtggaagaacctcacaaatgtgagttttgtggaaggagttacaagcaga 445
 DB 800 GGACGCACTCGTTGGTAAAGCTCACAATATGTTGGTATTTGTGGCCGGAGCTATAAACAGC 859
 QY 446 gaagttcccttgagagcagaagaagcgctgcctacattcttcagagcactgacc--- 503
 DB 860 GAAGCTTTTAGAGGAGCATAAAGAGCGATGCCACAACACTACTTTGGAAAGCATGGGCTTC 919
 QY 503 ---caggggacactgcagtgcgagggcagacacacaaag----- 542
 DB 920 CGGCGTGTGCCAGTCAATTAAGGAAGAACTAACCAACAGAGATGGCAGAACCTGT 979
 QY 542 cagagatgggaagtgaagagctctctactggacagattagcaagaatgtggcaaac 601
 DB 980 GCAAGATAGGAGCAGAGAGGTCCCTTCTCTCTGGACAGGCTGGCAAGCAATGTCCCAAC 1039
 QY 602 gaaaaagctcaatgcctcagaatt 626


```

QY 206 tggttcataagcgaagccatactagtggaacgccccattccagtgtaatacagtggtgggcat 265
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 TGGTTCACAAAAGAACCCACTGGAGAACGGCCCTTCAGTGCATCAGTGCAGGGCCT 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 266 cttttactcagaaggttaacctctccacattaaactgcacacaggggaaaaacctt 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 CATTCACCCAGAGGCAACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 326 ttaagtgcacctctgcaactatgcacaaagaagagatgcgtcacggggcatctta 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 TCAATAGCCACCTCTGCAACTACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 386 ggaacattctgtgagaacacctacaatgtgagtttgggaagagattacaagcaga 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 GGACGCACCTCCGTTGTAAACCTCAAAATGTGGATATTGTGGCCGAAGCTATAAACAGC 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 446 gaagttcccttgaggagcaagagcgctgcgtacatttcttcagagcactgacc--c 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 GAACGCTTTAGAGAACATTAAGAGCGCTGCCACAACACTACTTGGAAAGCATGGGCCCTTC 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 504 aggggacactgcaagtgcggagggaagacacacatcaagca----- 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 CGGGCACACTGTACCCAGTCATTAAAGAGAAACTAAGCACAGTGAATGGCAGAGACC 592
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 544 -----gagatgggaagtgaagagctctctgactggacagattagcaagaatgtggcaa 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 593 TGTGCAAGATAGGATCAGAGAGATCTCTGCTGTGACAGACTAGCAAGTAATGTGCGCA 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 599 aacgaaaaagctcaatgcctcagaatt 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 653 AAGCTAAGAGCTCTATGCCCTCAGAAATT 680
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 2

```

PCT-US93-08743-3
; Sequence 3, Application PC/TUS9308743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08743
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,233
; FILING DATE: 14-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1611 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1611
PCT-US93-08743-3

```

```

Query Match 34.5%; Score 216.4; DB 5; Length 1611;
Best Local Similarity 67.9%; Pred. No. 2.1e-61;
Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

QY 146 ccagtgaaagatgaactgcgctgtgtgattatctctgcacagctcaatgtcttaa 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 CTAACGGAAACTAAAGTGTGATCTGTGGGATCATTTGATCGGGCCCAATGTGCTCA 454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 206 tggttcataagcgaagccatactagtggaacgccccattccagtgtaatacagtggtgggcat 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 TGGTTCACAAAAGAACCCACTGGAGAACGGCCCTTCAGTGCATCAGTGCAGGGCCT 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 266 cttttactcagaaggttaacctctccacattaaactgcacacaggggaaaaacctt 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 CATTCACCCAGAGGCAACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 326 ttaagtgcacctctgcaactatgcacaaagaagagatgcgtcacggggcatctta 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 TCAATAGCCACCTCTGCAACTACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 386 ggaacattctgtgagaacacctacaatgtgagtttgggaagagattacaagcaga 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 635 GGACGCACCTCCGTTGTAAACCTCAAAATGTGGATATTGTGGCCGAAGCTATAAACAGC 694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 446 gaagttcccttgaggagcaagagcgctgcgtacatttcttcagagcactgacc--c 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 695 GAACGCTTTAGAGAACATTAAGAGCGCTGCCACAACACTACTTGGAAAGCATGGGCCCTTC 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 504 aggggacactgcaagtgcggagggaagacacacatcaagca----- 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 755 CGGGCACACTGTACCCAGTCATTAAAGAGAAACTAAGCACAGTGAATGGCAGAGACC 814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 544 -----gagatgggaagtgaagagctctctgactggacagattagcaagaatgtggcaa 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 815 TGTGCAAGATAGGATCAGAGAGATCTCTGCTGTGACAGACTAGCAAGTAATGTGCGCA 874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 599 aacgaaaaagctcaatgcctcagaatt 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 875 AAGCTAAGAGCTCTATGCCCTCAGAAATT 902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 3

```

PCT-US95-09345-2
; Sequence 2, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MGP-027PC
; TELECOMMUNICATION INFORMATION:

```

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 06:57:37 ; Search time 44.29 Seconds
(without alignments)
1475.260 Million cell updates/sec

Title: US-09-019-348-7
Perfect score: 628
Sequence: 1 gaaagagatgagaatgtttt.....ctcaatgctcagaattca 628

Scoring table: IDENTITY_NUC

Searched: 192659 seqs, 52021692 residues

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS9_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216.4	34.5	1386	3	US-08-465-590-3
2	216.4	34.5	1611	5	PCT-US93-08743-3
3	216.4	34.5	1386	5	PCT-US95-09345-2
4	209.8	33.4	2049	3	US-08-465-590-5
5	209.8	33.4	2049	5	PCT-US95-09345-4
6	170	27.1	1788	3	US-08-465-590-2
7	170	27.1	1788	5	PCT-US93-08743-2
8	170	27.1	1788	5	PCT-US95-09345-1
9	148.6	23.7	1296	3	US-08-465-590-4
10	148.6	23.7	1296	5	PCT-US95-09345-3
11	107.6	17.1	1170	3	US-08-465-590-6
12	107.6	17.1	1170	5	PCT-US95-09345-5
13	106.6	17.0	1004	3	US-08-465-590-8
14	106.6	17.0	1004	5	PCT-US93-09345-7
15	72.8	11.6	2031	4	US-08-933-750C-53
16	69.6	11.1	1309	4	US-08-933-750C-63
17	64.8	10.3	1892	4	US-08-933-750C-66
18	62	9.9	2133	3	US-08-820-170A-11
19	62	9.9	3754	3	US-08-820-170A-12
20	58.4	9.3	4252	5	PCT-US95-08429-4
21	56.6	9.0	3720	1	US-08-074-967-1
22	56.6	9.0	3720	4	US-08-553-541B-1
23	56.6	9.0	3720	5	PCT-US94-05669-1
24	55.2	8.8	3810	3	PCT-US95-08429-8
25	49.2	7.8	267	3	US-08-040-548-14
26	49.2	7.8	3086	3	US-08-040-548-15
27	49.2	7.8	267	3	US-08-466-344-14
28	49.2	7.8	3086	3	US-08-466-344-15
29	49.2	7.8	3068	3	US-08-224-482-1
30	48	7.6	2043	4	US-08-388-590A-39
31	48	7.6	5648	5	PCT-US96-03940-1
32	48	7.6	756	5	PCT-US96-03940-2
33	48	7.6	1407	5	PCT-US96-03940-3
34	48	7.6	1090	5	PCT-US96-03940-4
35	48	7.6	928	5	PCT-US96-03940-5
36	48	7.6	1791	5	PCT-US96-03940-6
37	48	7.6	3291	5	PCT-US96-03940-10

ALIGNMENTS

RESULT 1
US-08-465-590-3
: Sequence 3, Application US/08465590
: Patent No. 5824770
: GENERAL INFORMATION:
: APPLICANT: Georgopoulos, Katia A.
: TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
: NUMBER OF SEQUENCES: 164
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 STATE STREET, Suite 510
: CITY: BOSTON
: STATE: MASSACHUSETTS
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII (text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/465,590
: FILING DATE: 05-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/238,212
: FILING DATE: 02-MAY-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/121,438
: FILING DATE: 14-SEP-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/946,233
: FILING DATE: 14-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Myers, Paul L.
: REGISTRATION NUMBER: 35,695
: REFERENCE/DOCKET NUMBER: MPG-006C2DV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1386 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1386
US-08-465-590-3

Query Match 34.5%; Score 216.4; DB 3; Length 1386;
Best Local Similarity 67.9%; Pred. No. 1.9e-61;
Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

Qy 146 ccagtggaaagatggaactgcatgtgtgtggattatccctgcatcagcttcaatgtcttaa 205
Db 173 CTACGGAAACAAAGTGTGATATCTGTGGGATCATTTGTCATCGGGCCCAATGTGCTCA 232

KW *lxr*α⁰; transgene; transgenic animal; transgenic mouse; mIK-3;
 KW immunocompromised; immune system disorder; nervous system disorder;
 KW animal model; ss.
 OS Mus musculus.

QY 409 tacaatatgagtttttgggaagagattacaagcagagaagttcccttggagagacacaag 468
 Db 562 CACAATAATGTGATATTGTGGCGGAGCTATAAACAGCGAAGCTCTTTAGAGAGGCATAAA 621
 QY 469 gagcgtccgtacatttcttcagagcactgacc-----caggggacactgcaagtgcg 522
 Db 622 GAGCGATGCCACACTACTTGGAAAGCATGGCCCTTCGGCGGCTGTGCCAGTCAATTAAG 681
 QY 523 gaggaagacacacatacaag-----cagagatggggaagtgaagagcgt 564
 Db 682 GAAGAACTAACCAACAGAGATGSCAGAGACCTGTGCAAGATAGGAGAGAGGTCC 741
 QY 565 ctgctactggcagattagcagaatgtgcaaaaacgaaagcctcagagaa 624
 Db 742 CTGTGCTGGAGAGGCTGCGCAAGCAATGCGCAACGCTATGAGAGCTCTATGCTCAGAAA 801
 QY 625 tt 626
 Db 802 TT 803

RESULT 13
 V6968
 ID V6968 standard; cDNA; 1788 BP.
 AC V6968;
 DT 14-JAN-1999 (first entry)
 DE Murine Ikaros encoding cDNA mik-2.
 KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
 KW differentiation marker; immune system; corpus striatum; AIDS;
 KW Alzheimer's disease; ss.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT CDS 223..1518
 FT /tag= a
 FT /product= "mik-2"
 FT /trans_except= (pos:385..387,aa:Gln)
 PN US5824770-A.
 PD 20-OCT-1998.
 PF 05-JUN-1995; 463590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-582621/49.
 DR P-PSDB; W72671.
 PT Ikaros poly:peptide(s) - useful for treating disorders of immune system or corpus striatum
 PS disclosure: Column 51-56; 11pp; English.
 CC The present invention describes a purified peptide having at least one of the following properties: (a) it stimulates transcription of a DNA sequence under the control of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (c) it competitively inhibits the binding of a naturally occurring Ikaros isoform to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (d) it competitively inhibits Ikaros binding to Ikaros responsive elements; or (e) it inhibits protein-protein interactions of transcriptional complexes formed with naturally occurring Ikaros isoforms. The proteins, provided that they stimulate gene transcription under the control of delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit binding of naturally occurring Ikaros isoforms to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit Ikaros binding to Ikaros-responsive elements and/or inhibit protein-protein interactions of transcriptional complexes with naturally occurring Ikaros isoforms, can be used to treat immune system disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g. Alzheimer's disease. The present sequence encodes a specifically claimed mouse Ikaros protein.
 CC Sequence 1788 BP; 437 A; 486 C; 495 G; 370 T;

Query Match 27.1%; Score 170; DB 1; Length 1788;
 Best Local Similarity 67.1%; Pred. No. 3.8e-45;
 Matches 283; Conservative 0; Mismatches 115; Indels 24; Gaps 2;
 QY 229 ggtgaagcccccattccagtgtaatcagtggtgggacatcttttactcagaagaagtaacctc 288
 Db 382 GGTGAAGGGCTTTCCAGTGCACACAGCTGTGGGCGCTCTTTACCCAGAAAGGCAACCTC 441
 QY 289 ctgcgcacacattaaactgcacacaggggaaaaaccccttttaagtgtcacctctgcaactat 348
 Db 442 CTGCGGCACATCAAGTCGCACCTCGGGTGAGAGCCCTTCAAAATGCCATCTTTGCAACTAT 501
 QY 349 gcatgccaaaagaagagatgcgtcacggggcatcttaggcacacattctgtgggaaaacccc 408
 Db 502 GCCTGCCCGCGGAGGAGCGCCCTCACGGCCACCTGAGGAGCAGCACCCTGTTGGTAAGCCT 561
 QY 409 tacaatatgagtttttgggaagagttacaagcagagaagttcccttggagagacacaag 468
 Db 562 CACAATAATGTGATATTGTGGCGGAGCTATAAACAGCGAAGCTCTTTAGAGAGGCATAAA 621
 QY 469 gagcgtccgtacatttcttcagagcactgacc-----caggggacactgcaagtgcg 522
 Db 622 GAGCGATGCCACACTACTTGGAAAGCATGGCCCTTCGGCGGCTGTGCCAGTCAATTAAG 681
 QY 523 gaggaagacacacatacaag-----cagagatggggaagtgaagagcgt 564
 Db 682 GAAGAACTAACCAACAGAGATGSCAGAGACCTGTGCAAGATAGGAGAGAGGTCC 741
 QY 565 ctgctactggcagattagcagaatgtgcaaaaacgaaagcctcagagaa 624
 Db 742 CTGTGCTGGAGAGGCTGCGCAAGCAATGCGCAACGCTATGAGAGCTCTATGCTCAGAAA 801
 QY 625 tt 626
 Db 802 TT 803

RESULT 14
 V42805
 ID V42805 standard; cDNA; 1788 BP.
 AC V42805;
 DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform mik-2 cDNA.
 KW Ikaros; mik-2; transcription factor; mouse; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis; ss.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT CDS 223..1518
 FT /tag= a
 FT /trans_except= (pos:385..387, aa:Gln)
 PN CA2194256-A.
 PD 05-MAR-1998.
 PR 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR P-PSDB; W70963.
 PT New nucleic acid encoding Ikaros protein involved in early differentiation of lymphocytes - existing in several isoforms, and PT related products, used to treat e.g. immune diseases or cancer and PT to control cell differentiation
 PS Claim 7; Page 68-70; 158pp; English.
 CC This is the nucleotide of mouse Ikaros cDNA (isoform mik-2) that CC codes for a 431-amino acid zinc finger protein (see W70963) that is CC involved in the early differentiation of lymphocytes. A cDNA CC library constructed from the mature murine T cell line E14 was CC screened with a multimerised oligonucleotide (see V42829) derived CC from a protein binding site (see V42804) of the CD3-delta enhancer CC to identify T cell specific sequences that bind and mediate

CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
 CC inhibit protein-protein interactions of transcriptional complexes with
 CC naturally occurring Ikaros isoforms, can be used to treat immune system
 CC disorders, e.g. leukemia or AIDS, or corpus striatum disorders, e.g.
 CC Alzheimer's disease. The present sequence encodes a specifically
 CC claimed mouse Ikaros protein.
 SQ Sequence 2049 BP; 514 A; 531 C; 574 G; 430 T;

Query Match 33.4%; Score 209.8; DB 1; Length 2049;
 Best Local Similarity 67.1%; Pred. No. 5.7e-58;
 Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;

Qy 146 ccagtggaagatgaactgcgatgtgtgtgattatctctgcatacagttcaatgtcttaa 205
 Db 560 CTAACGGAAAACTAAAGTGTGATATCTCTGGGATCGTTTGGCCCGGCCCAATGTGCTCA 619

Qy 206 tgggttcataagcgaagccatactgtgaacccattccagttgaatcagttgtgggcat 265
 Db 620 TGGTTTCAAAAGAAAGTCACTGTGTGAACCGCTTTCCAGTGCACAGTCTGGGGCCT 679

Qy 266 cttttactcagaagaagttaacctctcccgccacattaaactgcacacaggggaaaaacctt 325
 Db 680 CTTTACCAGAAAGCAACCTCTCGCGGCACATCAAGCTGCAGTCGGTGAGAGCCCT 739

Qy 326 ttaagtgtcaactctgcactatgcatacgaagaagagatgcgtcagcgggcatctta 385
 Db 740 TCAAAATGCCATCTTTGCAACTATGCTCGCGGGAGGAGCGCCCTCACCGGCCACCTGA 799

Qy 386 ggacacattctgtgagaacccctacaatgtgattgttggaaggattacaagcaga 445
 Db 800 GGACGCACCTCGTGGTAAGCCTCAAAATGTGGATATTGTGGCGGAGCCTATAAACAGC 859

Qy 446 gaagttccttgaggagcacaagggagcgcgtacatttcttcagagcactgacc--- 503
 Db 860 GAAGCTCTTAGAGAGAGATGAAGAGCCATGCCACAACTACTTGGAAAGCATGGCCCTTC 919

Qy 503 ---caggggacactgcaagtgcggaggcgaacacacatcaaaag----- 542
 Db 920 CGGGCGTGTGCCAGTCACTTAAGGAAGAAACTAACCAACAGAGATGGCAGAACCTGT 979

Qy 542 cagagatgggaagtgaagaactctctgactggacagattagcaagaatgtggcaaac 601
 Db 980 GCAAGATAGGAGCAGAGAGTGCCTTGTCTGTGGACAGGCTGGCAGCAATGTGCGCAAAAC 1039

Qy 602 gaaaagctcaatgcctcagaatt 626
 Db 1040 GTAAGAGCTCTATGCTCAGAAATT 1064

RESULT 10
 V42808
 ID V42808 standard; cDNA; 2049 BP.
 AC V42808;
 DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform mik-1 cDNA.
 KW Ikaros; mik-1; transcription factor; mouse; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis; ss.
 OS Mus sp.
 FR Key Location/Qualifiers
 FT CDS 223..1779
 FT exon /*tag= a
 FT exon 223..384
 FT exon /*tag= b
 FT exon /*number= Ex1/2
 FT exon 385..643
 FT exon /*tag= c
 FT exon /*number= Ex3
 FT exon 644..810
 FT exon /*tag= d
 FT exon /*number= Ex4
 FT exon 811..933

FT /*tag= e
 FT /number= Ex5
 FT 934..1076
 FT /*tag= f
 FT /number= Ex6
 FT 1077..1779
 FT /*tag= g
 FT /number= Ex7
 PN CA2194256-A.
 PD 05-MAR-1998.
 PF 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (SEHO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR P-PSDB; W70966.
 PT New nucleic acid encoding Ikaros protein involved in early
 PT differentiation of lymphocytes - existing in several isoforms, and
 PT related products, used to treat e.g. immune diseases or cancer, and
 PT to control cell differentiation
 PS Claim 7; Page 75-77; 158pp; English.
 CC This is the nucleotide of mouse Ikaros cDNA (isoform mik-1) that
 CC codes for a 518-amino acid zinc finger protein (see W70966) that is
 CC involved in the early differentiation of lymphocytes. mik-1 cDNA
 CC was isolated from a mature murine T cell line E14 library using a
 CC 300 bp fragment from the 3' end of mik-2 cDNA (see V42805) as
 CC probe. 5 Different isoforms of mouse Ikaros (see V42805 and
 CC V42807-10) have been identified. These arise by differential
 CC splicing of Ikaros gene transcripts. Isoform mik-1 contains all 7
 CC exons. It is abundantly expressed in the early foetal liver, the
 CC maturing thymus and the postnatal spleen. The Ikaros gene is
 CC located at the proximal arm of murine chromosome 11. Ikaros
 CC proteins are suggested to play a role as a genetic switch
 CC (see V42806, V42811 and V42840) Ikaros sequences are very similar.
 CC The invention provides Ikaros nucleic acids, vectors and host cells
 CC expressing Ikaros proteins. These are used to treat T and B cell
 CC diseases (e.g. immune deficiencies caused by drugs, radiation or
 CC cancers), to control expression of heterologous genes placed under
 CC control of an Ikaros-responsive element, to treat nervous system
 CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
 CC amplification or differentiation, especially in haematopoietic
 CC cells. Some Ikaros isoforms are antagonistic of others and may be
 CC used to inhibit interaction with DNA sequences. The same effect
 CC can be achieved with Ikaros-binding oligonucleotides. Examining
 CC the expression of the Ikaros gene, or its allelic structure, can be
 CC used to assess risk of acquiring the above diseases.
 SQ Sequence 2049 BP; 514 A; 531 C; 574 G; 430 T;

Query Match 33.4%; Score 209.8; DB 1; Length 2049;
 Best Local Similarity 67.1%; Pred. No. 5.7e-58;
 Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;

Qy 146 ccagtggaagatgaactgcgatgtgtgtgattatctctgcatacagttcaatgtcttaa 205
 Db 560 CTAACGGAAAACTAAAGTGTGATATCTCTGGGATCGTTTGGCCCGGCCCAATGTGCTCA 619

Qy 206 tgggttcataagcgaagccatactgtgaacccattccagttgaatcagttgtgggcat 265
 Db 620 TGGTTTCAAAAGAAAGTCACTGTGTGAACCGCTTTCCAGTGCACAGTCTGGGGCCT 679

Qy 266 cttttactcagaagaagttaacctctcccgccacattaaactgcacacaggggaaaaacctt 325
 Db 680 CTTTACCAGAAAGCAACCTCTCGCGGCACATCAAGCTGCAGTCGGTGAGAGCCCT 739

Qy 326 ttaagtgtcaactctgcactatgcatacgaagaagagatgcgtcagcgggcatctta 385
 Db 740 TCAAAATGCCATCTTTGCAACTATGCTCGCGGGAGGAGCGCCCTCACCGGCCACCTGA 799

Qy 386 ggacacattctgtgagaacccctacaatgtgattgttggaaggattacaagcaga 445
 Db 800 GGACGCACCTCGTGGTAAGCCTCAAAATGTGGATATTGTGGCGGAGCCTATAAACAGC 859

FT CDS 1. .1386
 FT /*tag= a
 FT /product= "Ikaros"
 PD US5824770-A.
 PD 20-OCT-1998.
 PR 05-JUN-1995; 465590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GENO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 98-582621/49.
 DR P-PSDB: W72872.
 PT Ikaros poly:peptide(s) - useful for treating disorders of immune
 PT system or corpus striatum
 PS Disclosure; Column 55-58; ilipp; English.
 CC The present invention describes a purified peptide having at least one
 CC of the following properties: (a) it stimulates transcription of a DNA
 CC sequence under the control of a delta A element, an NFkB element or an
 CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
 CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
 CC consensus sequence; (c) it competitively inhibits the binding of a
 CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
 CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
 CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
 CC (e) it inhibits protein-protein interactions of transcriptional complexes
 CC formed with naturally occurring Ikaros isoforms. The proteins, provided
 CC that they stimulate gene transcription under the control of delta A
 CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit binding of naturally occurring Ikaros isoforms to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
 CC inhibit protein-protein interactions of transcriptional complexes with
 CC naturally occurring Ikaros isoforms, can be used to treat immune system
 CC disorders, e.g. leukemia or AIDS, or corpus striatum disorders, e.g.
 CC Alzheimer's disease. The present sequence encodes a specifically
 CC claimed human Ikaros protein.
 SQ Sequence 1386 BP; 350 A; 403 C; 395 G; 238 T;

Query Match 34.5%; Score 216.4; DB 1; Length 1386;
 Best Local Similarity 67.9%; Pred. No. 3.5e-60;
 Matches 34; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

Qy 146 ccagtgaagatgaactgcgattgtgtgattatctctgcatcagcttcaatgtcttaa 205
 Db 173 CTAACGGAARAACTAAAGTGTATCTCTGGGATCATTTGGATCGGGCCCAATGTCTCA 232
 Qy 206 tgggtcataagcgaagccatactggtgaacgccattccagtgtaactgagtggtggcgt 265
 Db 233 TGGTTCACAAAGAAGGACACACTGGAGAACGGCCCTTCCAGTGCATCAGTGCAGTGGCGGCT 292
 Qy 266 ctttactcagaaggaagtaacctctccgacacattaaactcacacaggggaaacatt 325
 Db 293 CATTACCCAGGAGGCAACCTGCTCCGGCACATCAAGTGCATTCGCGGAGAGGCCCT 352
 Qy 326 ttaagtgtcactctgcaactatcatgccaagaagagatgcgtcacgggggcatctta 385
 Db 353 TCAATGCCACCTCTGCAACTACGCCCTGCCCGGAGGGAGGCCCTCAGTGGCCACCTGA 412
 Qy 386 ggacacattctgtggagaacccctacaattgtgagttttgtggaagaggttacagcaga 445
 Db 413 GGAGCGCACTCGTGGTAAAGCCTCACAATGTGGATATTGTGGCGGAAGCTATAAAGAGC 472
 Qy 446 gaagttcccttgagagacagaggagcgtccgtacattcttcttcagagcactacc-c 503
 Db 473 GAACGCTTTTAGAGAACATAAGAGCGCTGCCACAACTACTTGGAAAGCATGGCCCTTC 532
 Qy 504 aggggacactgcaagtgcgggggaagacacacatcaaaagca----- 544
 Db 533 CGGGCACACTGTACCCAGTCTAATTAAGAGAGAACTAAGACAGTGAATGGCAGAGACC 592

Qy 544 -----gagatgggaagtgaagagctctctgactgacagattagcaagaatgtggcaa 598
 Db 593 TGTGCAAGATAGGATCAGAGAGATCTCTGCTGTGCACAGACTAGCAAGTAATGTGCCCA 652
 Qy 599 aacgaaaaagcctaactgcctcagaatt 626
 Db 653 ACGTAAGAGCTCTATGCTCTCAGAAAT 680
 RESULT 6
 V42806
 ID V42806 standard; cDNA; 1386 BP.
 AC V42806;
 DT 11-JAN-1999 (first entry)
 DE Human Ikaros isoform hik-1 cDNA.
 KW Ikaros; hik-1; transcription factor; human; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 1. .255
 FT /*tag= a
 FT /number= Ex3
 FT FT 256. .423
 FT /*tag= b
 FT FT /number= Ex4
 FT FT 424. .549
 FT /*tag= c
 FT FT /number= Ex5
 FT FT 550. .684
 FT /*tag= d
 FT FT /number= Ex6
 FT FT 685. .1386
 FT /*tag= e
 FT FT /number= Ex7
 PN CA2194256-A.
 PD 05-MAR-1998.
 PF 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GENO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 98-378292/33.
 DR P-PSDB: W70964.
 PT New nucleic acid encoding Ikaros protein involved in early
 PT differentiation of lymphocytes - existing in several isoforms, and
 PT related products, used to treat e.g. immune diseases or cancer and
 PT to control cell differentiation
 PS Claim 7: Page 70-72; 158pp; English.
 CC This is the nucleotide of human Ikaros cDNA (isoform hik-1) that
 CC codes for a 461-amino acid zinc finger protein (see W70964) that is
 CC involved in the early differentiation of lymphocytes. It was
 CC isolated from a Jurkat T cell line library using mouse Ikaros exon
 CC 7 cDNA as probe. The Ikaros gene maps between p11.2-p13 on human
 CC chromosome 7. The human and murine Ikaros sequences (see V42805-11
 CC and V42840) are highly conserved. Differential splicing of Ikaros
 CC gene transcripts gives rise to different Ikaros protein isoforms.
 CC The invention provides Ikaros nucleic acids, vectors and host cells
 CC expressing Ikaros proteins. These are used to treat T and B cell
 CC diseases (e.g. immune deficiencies caused by drugs, radiation or
 CC cancers), to control expression of heterologous genes placed under
 CC control of an Ikaros-responsive element, to treat nervous system
 CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
 CC amplification or differentiation, especially in haematopoietic
 CC cells. Some Ikaros isoforms are antagonistic of others and may be
 CC used to inhibit interaction with DNA sequences. The same effect
 CC can be achieved with Ikaros-binding oligonucleotides. Examining
 CC the expression of the Ikaros gene, or its allelic structure, can be
 CC used to assess risk of acquiring the above diseases.
 SQ Sequence 1386 BP; 352 A; 403 C; 397 G; 234 T;

Query Match

34.5%; Score 216.4; DB 1; Length 1386;

```
Query Match      34.5%; Score 216.4; DB 1; Length 1611;
Best Local Similarity 67.9%; Pred. No. 3.7e-60;
Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

Qy 146 ccagtggaaagatgaactgcgatgtgtggtgattatccctgcacagcttcaatgtcttaa 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 CTAACGGAAACTAAAGTGTGATCTGTGGGATCATTTTGCATCGGGCCCAATGTGCTCA 454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 206 tggttcataagcgagccatactggtgaacgcccattccagtgtaatacagttggtggcat 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 TGGTTCACAAGAGACCCACATCGAGACGCGCTTCACATGCAATCAGTCGCGGGCCT 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 cttttactagaaggttaacctctccgcacattaaactgcacacaggggaaacacctt 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 CATTCACCCAGAAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGGGGAGAGCCCT 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 ttaagtgtcacctctgcaactatgcatgccaaagaagatcgctcacggggcatctta 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 TCAATATGCCACCTCTGCAACTACGCTGCCCGGAGGACGCCCTCACTGGCCACCTGA 634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 gacacattctgtgagaaacacctacaatgtgagtttgtggaaggagttacaagcaga 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 635 GACGCACTCCGTGTAAACCTCAAAATGTGGATATGTGGCGAAGCTATAAACAGC 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 gaagttcccttgagagcacaaggagcgctgcgtacattttcagagcaactgacc--c 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 695 GAACGTCTTTAGAGGAACAATAAGAGCGCTGCCACAACACTACTTGGAAAGCATGGGCCITC 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 504 aggggacactgcaagtgcggaagcgaagacacacataaagca----- 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 755 CGGGCACACTGTACCCAGTCATTAAGAGAAGAACTAAGCACAGTGAATGGCAGAGACC 814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 -----gagatgggaagtgaagagctctcgtactgacagattagcaagcaatgtggcaa 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 815 TGTGCAAGTAGATCAGAGAGATCTCTGCTGGAGAGACTACGACAGATATGTGCGCA 874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 599 aacgaaaagctcaatgcttcagaatt 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 875 AACGTAAGAGCTCTATGCTCAGAAATT 902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
ID T16060 standard; CDNA; 1386 BP.
AC T16060;
DT 09-MAY-1996 (first entry)
DE Human Ikaros cDNA h1k-1.
KW Ikaros; transgene; transgenic animal; transgenic mouse; h1k-1;
KW immunocomprised; immune system disorder; nervous system disorder;
KW animal model; ss.
OS Homo sapiens.
PH Key Location/Qualifiers
FT exon 1..255
FT /tag= a
FT /label= Exon-3
FT exon 256..423
FT /tag= b
FT /label= Exon-4
FT exon 424..549
FT /tag= c
FT /label= Exon-5
FT exon 550..681
FT /tag= d
FT /label= Exon-6
FT exon 682..1383
FT /tag= e
FT /label= Exon-7
PN WO9604372-A1.
PD 15-FEB-1996.
PF 28-JUL-1995; U09345.
PR 29-JUL-1994; US-283300.
PA (GEO ) GEN HOSPITAL CORP.
```

```
Georgopoulos K;
WPI; 96-129389/13.
P-PSDB; R92015.
Transgenic rodent having Ikaros trans-gene (pref. mutated) - is*
severely immuno-compromised and can be used as model to determine
effects of treatment for immune and nervous system disorders
PT Disclosure; Fig 2; 102pp; English.
CC An almost full-length cDNA sequence (T16060) codes for part
(R92015) of the human Ikaros protein, a master regulator of
CC haematopoietic differentiation. It was isolated from a T-cell
CC line Jurkat cDNA library using a mouse Ikaros cDNA clone as probe.
CC The human Ikaros gene maps to pl1.2-pl3 on chromosome 7.
CC Different isoforms (see R92014 and R92016-19) of mouse Ikaros have
CC also been isolated. Transgenic animals, pref. mice, having a
CC mutated Ikaros transgene, esp. a mutation that alters the DNA
CC binding domain of the Ikaros protein, are used as models to
CC determine the effects of treatments for immune or nervous system
CC disorders. 1386 BP; 350 A; 403 C; 395 G; 238 T;
SQ Sequence

Query Match      34.5%; Score 216.4; DB 1; Length 1386;
Best Local Similarity 67.9%; Pred. No. 3.5e-60;
Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

Qy 146 ccagtggaaagatgaactgcgatgtgtggtgattatccctgcacagcttcaatgtcttaa 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 CTAACGGAAACTAAAGTGTGATATCTGTGGGATCATTTTGCATCGGGCCCAATGTGCTCA 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 206 tggttcataagcgagccatactggtgaacgcccattccagtgtaatacagttggtggcat 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 TGGTTCACAAGAGACCCACATCGAGACGCGCTTCACATGCAATCAGTCGCGGGCCT 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 cttttactagaaggtgaacctctcgtccacattaaactgcacacaggggaaacacctt 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 CATTCACCCAGAAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGGGGAGAGCCCT 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 ttaagtgtcacctctgcaactatgcatgccaaagaagatcgctcacggggcatctta 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 TCAATATGCCACCTCTGCAACTACGCTGCCCGGAGGACGCCCTCACTGGCCACCTGA 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 gacacattctgtgagaaacacctacaatgtgagtttgtggaaggagttacaagcaga 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 GGACGCACTCCGTGTGTAACCTCACAAATGTGGATATTTGGCCGAAAGCTATAAACAGC 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 gaagttcccttgagagcacaagagcgctgcgtacattttcagagcaactgacc--c 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 GAACGTCTTTAGAGGAACAATAAGAGCGCTGCCACAACACTACTTGGAAAGCATGGGCCITC 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 504 aggggacactgcaagtgcggaagcgaagacacacataaagca----- 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 CGGGCACACTGTACCCAGTCATTAAGAGAAGAACTAAGCACAGTGAATGGCAGAGACC 592
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 -----gagatgggaagtgaagagctctcgtactgacagattagcaagcaatgtggcaa 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 593 TGTGCAAGTAGATCAGAGAGATCTCTGCTGGAGAGACTACGACAGATATGTGCGCA 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 599 aacgaaaagctcaatgcttcagaatt 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 653 AACGTAAGAGCTCTATGCTCAGAAATT 680
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
V66969
ID V66969 standard; CDNA; 1386 BP.
AC V66969;
DT 14-JAN-1999 (first entry)
DE Human Ikaros encoding cDNA.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; immune system; corpus striatum; AIDS;
KW Alzheimer's disease; ss.
OS Homo sapiens.
PH Key Location/Qualifiers
```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 02:38:40 ; Search time 70.34 Seconds
(without alignments)
2233.731 Million cell updates/sec

Title: US-09-019-348-7
Perfect score: 628
Sequence: 1 gaaagagatgagaatgtttt.....ctcaatgcctcagaattca 628

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	628	100.0	628	1 T60491	Human Aiolos parti
2	493.6	78.6	1984	1 T60490	Mouse Aiolos cDNA.
3	216.4	34.5	1611	1 Q44980	Human Ikaros pepti
4	216.4	34.5	1386	1 T16060	Human Ikaros cDNA
5	216.4	34.5	1386	1 V66969	Human Ikaros encod
6	216.4	34.5	1386	1 V42806	Human Ikaros isofo
7	216.4	34.5	1551	1 V42840	Human Ikaros isofo
8	209.8	33.4	2049	1 T16062	Murine Ikaros cDNA
9	209.8	33.4	2049	1 V66971	Mouse Ikaros encod
10	209.8	33.4	2049	1 V42808	Mouse Ikaros isofo
11	170	27.1	1788	1 Q44979	Murine Ikaros gene
12	170	27.1	1788	1 T16059	Murine Ikaros cDNA
13	170	27.1	1788	1 V66968	Murine Ikaros enco
14	170	27.1	1788	1 V42805	Mouse Ikaros isofo
15	148.6	23.7	1296	1 T16061	Murine Ikaros cDNA
16	148.6	23.7	1296	1 V66970	Mouse Ikaros encod
17	148.6	23.7	1296	1 V42807	Mouse Ikaros isofo
18	107.6	17.1	1170	1 T16063	Murine Ikaros cDNA
19	107.6	17.1	1170	1 V66972	Mouse Ikaros encod
20	107.6	17.1	1170	1 V42809	Mouse Ikaros isofo
21	106.6	17.0	1004	1 T16065	Mouse Ikaros cDNA. Trans
22	106.6	17.0	1004	1 V66974	Ikaros isoform enc
23	106.6	17.0	1004	1 V42811	Human Ikaros cDNA.
24	106.2	16.9	168	1 V67125	Ikaros protein enc
25	104.6	16.7	168	1 V67126	Ikaros protein enc
26	71.6	11.4	1663	1 X39662	Renal cancer assoc
27	68.4	10.9	323	1 T25378	Human gene signatu
28	66.4	10.6	765	1 X39735	Gastric cancer ass
29	64.6	10.3	851	1 X00648	Human secreted pro
30	63.8	10.2	345	1 T25341	Human gene signatu
31	63.8	10.2	353	1 V86582	EST clone AW95. Ne
32	62.8	10.0	398	1 Q59865	Human brain expres
33	62.8	10.0	1683	1 V73484	Human SRE-2BP anal
34	62.8	10.0	2168	1 V73485	Human SRE-2BP anal
35	62	9.9	2133	1 V01890	Human OTK18 gene.
36	62	9.9	3754	1 V01891	Human OTK18 gene.
37	61.6	9.8	264	1 V83634	DNA encoding a nuc
38	61.6	9.8	264	1 V99467	DNA encoding an an
39	61.6	9.8	264	1 X16976	Zinc finger protei
40	60.2	9.6	270	1 V89789	EST clone COL069.
41	59.8	9.5	2680	1 V64579	Myc-binding zinc-f
42	59.6	9.5	135	1 V67122	Ikaros protein enc
43	58.2	9.3	3099	1 V64361	Human stem cell zi

44 56.6 9.0 3720 1 Q80513 Genetic locus bcl-
45 56.4 9.0 3776 1 T42903 TRP-1 protein codi

ALIGNMENTS

RESULT 1					
T60491					
ID T60491 standard; cDNA; 628 BP.					
AC T60491:					
DT 07-JUL-1997 (first entry)					
DE Human Aiolos partial cDNA.					
KW Aiolos; transcription activator; immune system; T lymphocyte;					
B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;					
KW transgenic animal; ss.					
OS Homo sapiens					
TN W09714714.A1.					
PD 24-APR-1997.					
PF 17-OCT-1996; U16774.					
PR 18-OCT-1995; US-005529.					
PR 14-MAY-1996; US-017646.					
PA (GEHO) GEN HOSPITAL CORP.					
PI Georgopoulos K, Morgan BA;					
DR WPI; 97-245047/22.					
DR P-PSDB; W15575.					
PT Aiolos polypeptide and corresponding DNA - used to reconstitute a					
PT mammalian immune system, for the treatment of T cell leukaemia(s)					
PT and lymphoma(s)					
PS Disclosure; Fig 5a; 115pp; English.					
CC Partial human Aiolos cDNA (T60491) was isolated by PCR					
CC amplification using primers (see also T60494-95) based on mouse					
CC Aiolos gene exons. The Aiolos gene is a homologue of Ikaros whose					
CC expression is restricted to lymphoid lineage. Aiolos polypeptide					
CC can form dimers with Aiolos or Ikaros polypeptides, is expressed in					
CC committed lymphoid progenitors and in T and B cells, and is a					
CC transcriptional activator of a lymphoid gene. Aiolos cDNA can be					
CC used to produce recombinant Aiolos polypeptide (W15575) in host					
CC cells. The Aiolos polypeptide, coding sequence or cells expressing					
CC therapy. Such disorders include T-cell leukaemias and lymphomas.					
CC Non-wildtype gene structure or expression is indicative of a risk					
CC for such a disorder. Transgenic animals with an Aiolos transgene					
CC are provided.					
SQ Sequence 628 BP; 207 A; 128 C; 155 G; 138 T;					
Query Match 100.0%; Score 628; DB 1; Length 628;					
Best Local Similarity 100.0%; Pred. No. 2.5e-193;					
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy 1 gaaagagatgagaatgttttaagtcagaaacccatgggaatgcagaagagccctgaaatc 60					
Db 1 GAAAGAGATGAGATGTTTTTAAGTCAGAACCCCATGGGAATGCAGAGAGCCCTGAAATC 60					
Qy 61 cctcacagctattcaagaataataatgaataatgaaacatttaagtgtgagagacatgtt 120					
Db 61 CCTTACACCTATTCAAGAGAAATATAATGAATATGAATAATTAAGTTGGAGACATGTT 120					
Qy 121 gtctcaattcgtatgtagcaggcccaaccagtggaaagatgaactgcgagtgtgtggatta 180					
Db 121 GTCTCATTCGATAGTAGCAGGCGCAACCAAGTGGAAAGATGAACCTGCGATGTGTGGATTA 180					
Qy 181 tccgcacatcagctcaatgtcttaagtgttcataagcgaagccattctggtgaagcccca 240					
Db 181 TCCGTCATCAGCTCAATGCTTTAATGGTCTTAAGCGAAGCCATCTGGTGAAGCCCCA 240					
Qy 241 ttccagtgtaatcagtggtgggcatcttttactcagaagaagtaacctctccgccacatt 300					
Db 241 TTCCAGTGTAAATCAGTGTGGGCATCTTTTACTCAGAAAGGTAACTCTCCGCCACATT 300					
Qy 301 aaactgcacacaggggaaacaccttttaagtgtcacctctgcaactatgcatgccaaga 360					

56215 AATTACCCCTGTCCCGGAGAGAGCCCTGACCGGACATCTACGCACGCACGCAG 56269

Search completed: November 6, 1999, 06:39:51
Job time: 4705 sec

U92199.1	GI:2062737
VERSION	rainbow trout.
KEYWORDS	Oncorhynchus mykiss
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
ORGANISM	Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
REFERENCE	Salmoniformes; Salmonidae; Oncorhynchus.
AUTHORS	Hansen,J.D.
TITLE	Isolation and characterization of Ikaros homologues in the rainbow trout
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2079)
AUTHORS	Hansen,J.D.
TITLE	Direct Submission
JOURNAL	Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
FEATURES	Location/Qualifiers
source	1..2079
	/organism="Oncorhynchus mykiss"
	/strain="shasta"
	/db_xref="taxon:8022"
gene	/cell_type="thymocyte"
	81..1232
CDS	/gene="Ikaros"
	81..1232
	/note="Ik-8 isoform; similar to mouse and human Ikaros/LyF-1; alternatively spliced form missing exons 3 and 8"
	/codon_start=1
	/product="Ikaros homolog"
	/protein_id="AAB53433.1"
	/db_xref="PID:G2062738"
	/db_xref="GI:2062738"
	/translation="MEMEAEQMSOMPGRDPPNDVSEENDEAMPIDELSSANNLQHNNRGKEGLRPFCTCGASFTKGNLLRIHKLHSGERPKFCNLCNYACRRDALSGHLRTHSVKPCYAGCNSAYKQSLEEKERLYNQCLMGQNSYITGEKRFNSISFEFGPELMOPHVDOAINSAINYLGAESLRPLQTSTSDMGVMGSMTPLHKPAAEGHLSAKDSAENLLAKSKASSEKDGSPSHSGODSTOTENNEKEAGVGASGLIYLTNHTSGVRNGVLVKVEOROEYAMRASIEIASPGFKVLSCEGOVRYRCHECRILFLDHWNVYTHMCHGFDPDFECNLGHRSDRYEFSSHWTGHEHY"
BASE COUNT	593 a 500 c 464 g 521 t 1 others
ORIGIN	
Query Match	25.3%; Score 158.6; DB 4; Length 2079;
Best Local Similarity	59.6%; Pred.No.2e-35;
Matches	215; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy	229 ggtagagccattcaggtaaatcagtgtggggcatctttactcagaagaagtaacctc 288
Db	243 GGAGAGGTCCATCCAGTCACCAGTGTGGCGCTCTTTACCCAGAGAAGCAACTG 302
Qy	289 ctccgcacatataactgcacacaggggaaaaaccctttaagtgtcacctctgcaactat 348
Db	303 CTACGTACATCAAGTCCACTCAGGAGAGAACCCCTTAAGTGTCACCTGTCAACTAT 362
Qy	349 qcatgccaaagaagatcgcttcacaggggcatcttagacacatctcttgagaaaacc 408
Db	363 GCCTGTGCCGGAGAGACGCCCTCACGGGTCACCTGCATACCATCTGTGGAAAACCC 422
Qy	409 tacaaatgtgagtttgtgaagaggttacaagcagagaagttccccttgagagacaga 468
Db	423 CACAAGTGTCTTACTGTGGCGGAGCTACACAGCAGTAGTCTCTGTGAGGAACACAG 482
Qy	469 gagcgtgcctgatattcttcagacactgacccaggggacactgcaagtgcaggagca 528
Db	483 GAGCGGTGTACAACTACCTCCAGTGTGATGGGCTCCAGACAGCATATACAGGTGAG 542
Qy	529 agacacatc 537
Db	543 AAACGCTTC 551

RESULT	15		
AF056116	DNA	VRT	15-APR-1999
LOCUS	148640 bp		
DEFINITION	Fugu rubripes serine/threonine kinase receptor type1, All-1 related protein (ALR), fugu hedgehog (fhh), ikaros-like, wnt1, wnt10b, ARF3, erbB3, FAS1, and L41 ribosomal protein genes, complete cds; LRP1 gene, partial cds; and unknown genes.		
ACCESSION	AF056116 U82608		
NID	G3170534		
VERSION	AF056116.1 GI:3170534		
KEYWORDS	Fugu rubripes		
SOURCE	Fugu rubripes		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidae; Tetraodontidae; Fugu.		
REFERENCE	1 (bases 1 to 148640)		
AUTHORS	Gellner,K. and Brenner,S.		
TITLE	Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu rubripes		
JOURNAL	Genome Res. 9 (3), 251-258 (1999)		
MEDLINE	99177347		
REFERENCE	2 (bases 1 to 148640)		
AUTHORS	Gellner,K. and Brenner,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-1998) Molecular Sciences Institute, 2168 Shuttock Ave, Berkeley, CA 94704, USA		
COMMENT	On Jun 2, 1998 this sequence version replaced gi:1814290. Prediction of protein sequences are solely based on homology search analysis and XGRAIL analysis of the genomic DNA sequence.		
FEATURES	Location/Qualifiers		
source	1..148640		
	/organism="Fugu rubripes"		
	/db_xref="taxon:31033"		
	/clone="J9P14; J8; 138E3; 186G5; 178O23"		
	/note="sequenced from overlapping cosmids of the HGMP Fugu cosmid library"		
CDS	join(449..545,702..943,1038..1281,1393..1623,1735..1902,2006..2162,2229..2353,2432..2562,2649..2774) /codon_start=1 /product="serine/threonine kinase receptor type1" /protein_id="AAC34382.1" /db_xref="PID:g3170535" /db_xref="GI:3170535" /translation="MLQWRHGSRFLLLGLFIYTSHWASSVSVCALARCNCTCKEYCYCEDTGACMASTYYVERKQHIRICINRDLVPQPGFYCLSAEGLNTHCCYEY CNSIDLVTKCEGWASSMSGPVAVIAGVFLLCVLMGVLFQYHQRYASHRD RLEVEDPSDDHIYLPKDKTLQDIYDMSTSGSGSLPQVQRTVARTIVLOIIIGKR FGWVRGKWGGDVAVKIFSSREERSWFREAELYOTIMLRHENILGLFIAADNKDGTW TOLWLVDYHEGSLFDYLNRYSVTEGMKKALSASGLAHLHWEILLTGCKPGIAH RDLKSNIYKKNMCIAIDLGLAVRHESIITDITAPNRQVGTGYMAPVELDETIN MKHFDSFKCAIALGLVWEIARENCAGGIIHEEVOLPYDILVPSDPSIEEMRKVVCD QKLRPNVPMWQSYEALRYMKRECWYANGAARLTALRIKKTLSQLSVEEDIKM" complement(7874..29042) /gene="ALR"		
gene	complement(join(7874..7966,8133..8241,8314..8387,8531..8816,8931..9061,9157..9293,9381..10242,10375..10740,10821..10939,11300..11435,11512..11636,11724..11896,11977..12058,12137..12284,12369..12533,12619..12762,12819..15788,15881..16113,16195..16261,16348..16429,16726..16849,16931..18699,18902..19041,19230..19421,19507..21090,21210..21260,21606..21679,21759..21997,22095..22164,22287..22442,22670..22982,23076..23179,23295..23417,23512..23688,23787..24006,24177..24300,24454..24618,24844..25025,25108..25212,25295..25408,25539..25646,25735..26090,26209..26301,26745..27040,27126..27304,27408..27522,27613..27781,27884..28045,28160..28269,28364..28566,28677..28779,28994..		

```

CDS
<1..>415
/gene="Ikaros"
/notes="similar to mouse and human Ikaros/LyF-1; composed
of exons 4, 5 and 6"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53435.1"
/db_xref="PID:g2062744"
/db_xref="GI:2062744"
/translation="GERPFOCNCGASFTQKGNLLRHILHSGEKFCHMCNYACRR
RDALSGHLRTHSVGKPKHCYGRSKYKORSSLEHKEKCHNYLQSIGLQHLVAYKEE
SKTDMEDLSKIGESRLVLDRLASNAVKRKSMPQ"
BASE COUNT 124 a 93 c 100 g 98 t
ORIGIN

Query Match 26.5%; Score 166.2; DB 4; Length 415;
Best Local Similarity 65.3%; Pred. No. 1.4e-37;
Matches 27; Conservative 0; Mismatches 123; Indels 21; Gaps 1;

Qj 229 ggtgaacgcccattccagtgtaacagtggtgggcatcttttactcagaaggttaacctc 288
Nh 1 GCTGAACGGCCTTCCAGTGCACACAGTGTGTCCTCTTACCCAGAGGGCAATCTC 60

Qy 289 ctccgcccattaaactgcacacaggggaaaaaaccttttaagtacacctcgcaactat 348
Db 61 CTTCTCATATCAAAATACACTCTGCTGTAAGAGCCCTTTAAATGTCATGTGTAACAT 120

Qy 349 gcatgccaagaagagatgcgtccacgggcatcttaggcacacattctgtggaaacccc 408
Db 121 GCATGCGAGCGCGAGGATGCATCTAGTGGTCACCTCAGGACACATTTCTAGGCAACCT 180

Qy 409 tacaatgtgagttttgtggaaggagtacaagcagagaagttcccttgaggagcacaag 468
Db 181 CACAAGTGTGATACTGTGTCGTAGCTATAGCAACGGAGCTCTTAGAAGAGCATAAA 240

Qy 469 gaggcgtgcgtacattcttcagagacactgacccaggggacactgcaagtgcggagcca 528
Db 241 GAGAGATGTCAACAACCTACCTACAAGTATAGGCTTGCGAGGCCCATCTTTATGCTGTAAA 300

Qy 529 agacacatcaaaa-----gcagagatgggaagtgaagagagctctc 567
Db 301 GAAGAGACCAAAATGATATGGCAGACACCTGTCCAGATGGGTGAGAGATCACTT 360

Qy 568 gtactgacagattagcaagcaatgtggcaaaacgaaagctcaatgcctcaga 622
Db 361 GTGCTTGACAGACTGCAAGTAAGCTTGCCAAAGCTAAGAGCTCTATGCCTCAGA 415

RESULT 13
OMU92200
LOCUS
DEFINITION
Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, 02-MAY-1997,
complete cds.
ACCESSION
U92200
NID
g2062739
VERSION
U92200.1
KEYWORDS
rainbow trout.
SOURCE
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
Hansen, J.D.
Isolation and characterization of Ikaros homologues in the rainbow
trout
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2183)
AUTHORS
Hansen, J.D.
TITLE
Direct Submission
JOURNAL
Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland

```

```

FEATURES
Source
Location/Qualifiers
1..2183
/organism="Oncorhynchus mykiss"
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
50..1336
/gene="Ikaros"
50..1336
/gene="Ikaros"
/notes="Ik-2 isoform: similar to mouse and human
Ikaros/LyF-1; alternatively spliced form missing exon 3"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53434.1"
/db_xref="PID:g2062740"
/db_xref="GI:2062740"
/translation="MEMEPAOEMSOMPGRDSPPNDVSEENDEAMPIDEDLSASSNLQ
HNRRGKEGLGERPFOCTCGASFTQKGNLLRHILHSGEKFCHLCNYACRRRAL
SGHLRTHSVGKPKHCAYCGRSKYKORSSLEHKEKCHNYLOCMLQNSIYTVKEESNQ
NEQREDLSQMSKRALVLDRLANNVAKRSTMPQKVFGEKRFNSIFEGGPELMOPH
VIDQAINSAINYLGAEISLPLQTSPTSSDMGVMGSMYPLHPPAPAGHLSKDSAAE
NLLLLAKSKASSEKDGSPSHGQSTDTSENNEKAGVAGSLIYLTNHTSGVRNG
VLPLVKEEOQROYEAMRASIEIASIEGKSVLSGEQVRAVRCEHCRIFLDHWMTIHI
MCHGFRDPFECLNCGHRSQDRYERSSHMTREHRY"
BASE COUNT 638 a 517 c 488 g 539 t 1 others
ORIGIN

Query Match 26.4%; Score 165.8; DB 4; Length 2183;
Best Local Similarity 64.9%; Pred. No. 1.8e-37;
Matches 276; Conservative 0; Mismatches 122; Indels 27; Gaps 1;

Qy 229 ggtgaacgcccattccagtgtaacagtggtgggcatcttttactcagaaggttaacctc 288
Db 212 GGAGAGCGTCCATTCAGTGCACCCAGTGTGGCGTCTTTTACCCAGAGGCAACCTG 271

Qy 289 ctccgcccattaaactgcacacaggggaaaaaaccttttaagtacacctcgcaactat 348
Db 272 CTACGTCAACATCAAGTCTCCACTCAGGAGAGAGCCCTTCAAGTGTCACTGTGCAACTAT 331

Qy 349 gcatgccaagaagagatgcgtccacgggcatcttaggcacacattctgtggaaacccc 408
Db 332 GCCTGTGCGGAGAGAGCGCCCTCAGCGGTACCTGCGGTACCCACTCTGTGTGAAAACCC 391

Qy 409 tacaatgtgagttttgtggaaggagtacaagcagagaagttcccttgaggagcacaag 468
Db 392 CACAAGTGTGCTTACTGTGGCGGAGCTACAGCAGCGTAGCTCTCTGTGGAGAACACAAG 451

Qy 469 gaggcgtgcgtacattcttcagagacactgacccaggggacactgcaagtgc----- 522
Db 452 GAGCGGTGTCAACAACCTACCTCCAGTGCATGGGCTCCAGAACACAGCATCTATACAGTAGTA 511

Qy 522 -----ggaggcaagacacatcaaacgacagatggaagtgaagaaga 561
Db 512 AAGGAAGAAAGCAACCAAGATGAGCAGAGGGAAGACTTAAAGCCAGATGGGATCTTAAGAGA 571

Qy 562 gctctcgtagtgacagattagcaagcaatgtggcaaaacgaaagctcaatgcctcag 621
Db 572 GCCTTGTGCTAGACACACTAGCTAATAATGTAGCCAAACGTAAGACACTATGCCACAG 631

Qy 622 aaatt 626
Db 632 AAGTT 636

RESULT 14
OMU92199
LOCUS
DEFINITION
Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, 02-MAY-1997,
complete cds.
ACCESSION
U92199
NID
g2062737

```



```
Oy 386 ggacacattctgtgagaaccctacaataatgtgagttttgtggaaggagttacaagcaga 445
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 578 GGACTCACTCGTGGTGGCAACCCCAATAAGTGTGGTACTGTGGTGCAGCTATAAGCAGC 637
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 446 gaagttcccttgaggagcacaaggagcgtccgtacatttttcagagca--ctgaacc 503
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 638 CGAGCTCTTGGAGAACATAAAGAACGCTGTGCATAACTACTGCAACCATGAGTACT 697
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 504 aggggacactgcaagtgcggagggcagaagcacatcaaacag----- 545
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 698 CAAGCAATCTTTATTCAGTCAATAAAGAGAGAACTAACAGAGTGAATGGCTGAAGACC 757
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 545 -----agatgggaagtgaagagctctgtactggacagattagcagaatgtggcaa 598
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 758 TGTGCAAGATAGGGTCAGAAAGATCCCTGTGCTGGATAGACTAGCAAGTAACGTCGCCA 817
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 599 aacgaaaaagctcaatgcctcagaatt 626
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 818 AACGTAAGAGCTCTATGCTCAGAAATT 845
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
OMU92198 2301 bp mRNA VRT 02-MAY-1997
LOCUS Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, Ik-7 isoform,
DEFINITION complete cds.
ACCESSION U92198
NID 92062735
VERSION U92198.1 GI:2062735
KEYWORDS rainbow trout.
SOURCE Oncorhynchus mykiss
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 2301)
AUTHORS Hansen, J.D.
TITLE Isolation and characterization of Ikaros homologues in the rainbow
trout
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2301)
AUTHORS Hansen, J.D.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
FEATURES
source 1. 2301
/organism="Oncorhynchus mykiss"
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
36. 1469
/gene="Ikaros"
36. 1469
/gene="Ikaros"
/translation="MEMEAAQMSQMPGRDPPPPNDLSEENDEAMPIDLSASSNLQ
HNRRDGLACNTKVEARCDENGLAIDMMNGEEDDECAEDLRVLDASCAKVGSH
AGGDPKSGYPSAGGIRLPNGKLKDCIGVICIPNVLVHVKRSHGTERPPOCTQCGA
SFTOKNLLRHIKLHSGEKFPEKHLNACVRRDALSGHLFTHSVGPHKCAYGRSY
KQRLSEHKRCHNYLQCMGLQNSITGKERFNSFEFGPGELMQPHVLDQAINSA
INYLGAESLRPIQTSPTSDMGVNGSMYPLHKPPAEGHLSAKDSAAENLLAKSK
SASKEKQSPSHSQDSDTDESNNEEKAGVSLIYLNTHITSGVRNGVLLVKEEQ
QRYEAMRASIEIAEGFKVLSGEQVRAVRCHECRILFLDHDHMYITHMCHGFRDP
FECNLGHRSDRVEFFSHMTGEGHY"
BASE COUNT 625 a 562 c 545 g 568 t 1 others
ORIGIN
```

```
Query Match 32.5%; Score 204; DB 4; Length 2301;
Best Local Similarity 69.1%; Pred. No. 1.9e-48;
Matches 279; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Oy 134 gtacagagcccaacagtggaagatgaactcgcgtgtgtgtgattatcctgcatacgt 193
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 GTATTGCGCTGCCCAACGGGAAGCTCAAGTGTGATATCTGTGGGATAGTGTTCATTGGCC 444
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 194 tcaatgtcttaatggttcataagcgaaacactactggggaacgcccatcctcagtgtaact 253
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CCAATGTCTGATGTGCAAGCGAAGCAACACACTGGAGCGGTCCATTCAGTGCACCC 504
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 254 agtgtggggcctcttttactcagaagagtaaacctctctccgcacattaaactgcacacag 313
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 AGTGTGGCGCTTCTTTCACCCAGAAAGCAACCTGCTACGTACATCAAGCTCCACTCAG 564
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 314 gggaaaaaaccttttaagtgtcacctctgcactatgcataaagaagagatgcgtca 373
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 GAGAGAAAGCCCTTCAAGTGTCACTGTGCAACTATGCTGTCCGCGGAGAGACGCCCTCA 624
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 374 cggggcattctagacacattctgtgagaacccctacaataatgtgagttgtggaaga 433
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 CGGTCACTCGGTACCCACCTCTCTGTGGAAACCCCAAGTGTCTTACTGTGGCGGA 684
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 434 gttacaagcagagaagttcccttgaggagcacaaggagcgtgcgtacattttcttcaga 493
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GCTACAAGCAGCGTAGCTCTCTGGAGGAACACAGGAGCGGTGTCCAACTACCTCCAGT 744
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 494 gcactgacccaggggagacactgcaagtgcggagggcaagacacatc 537
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 GCATGGGGCTCCAGAACAGCATCTATACAGGTGAGAAACGCTTC 788
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
S74708 684 bp mRNA ROD 12-MAY-1995
LOCUS Ikaros/LyF-1-lymphoid transcription factor LyF-1 isoform VI
DEFINITION [alternatively spliced] [mice, RLM1, Tdt+ thymoma, mRNA Partial,
684 nt].
ACCESSION S74708
NID 9807152
VERSION S74708.1 GI:807152
KEYWORDS
SOURCE Mus sp. RLM1 Tdt+ thymoma.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 684)
AUTHORS Hahn, K., Ernst, P., Lo, K., Kim, G.S., Turck, C. and Smale, S.T.
TITLE The lymphoid transcription factor LyF-1 is encoded by specific,
alternatively spliced mRNAs derived from the Ikaros gene
JOURNAL Mol. Cell. Biol. 14 (11), 7111-7123 (1994)
MEDLINE 95021239
REMARK
This sequence comes from Fig. 7A.
This sequence comes from the original journal article.
FEATURES
source 1. 684
/organism="Mus sp."
/db_xref="taxon:10095"
4. 684
/partial
/gene="Ikaros/LyF-1"
/translation="lymphoid transcription factor LyF-1 isoform VI"
4. 684
/partial
/gene="Ikaros/LyF-1"
/translation="This sequence comes from Fig. 7A."
/codon_start=1
/exception="Protein longer than coding region shown"
/product="lymphoid transcription factor LyF-1 isoform VI"
```

```
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
1..2496
/gene="Ikaros"
81..1649
CDS
/gene="Ikaros"
/feature="Ikaros"
/note="Ik-1 isoform; similar to mouse and human
Ikaros/Lyf-1; Ikaros/Lyf-1 homolog"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AA97474.1"
/db_xref="PID:g2062742"
/db_xref="GI:2062742"
/translation="MEMEAQEMSQMPGRDPPNDVSEENDEAMPIDELASNLQ
HNRRDGLACINIKVEARCEENGLAIDMMNGEEDLRLVDSAGKAVNGSH
AGPQKGYPSAGGIRLPNGKLDICIGVICIPNVLVHNRSHTRGPFQCGTQCGA
SFTQGNLRLHKLHSGEPKCHLCNYACRRDALSHLRTSHVSKPHKAYCGRSY
KORSILEHKECHNYLOCMGLONSTYTVVKEESNONEREDLSOMGSKRALVLDLA
NVVAKKSTMPKQFVGEKRFNSISEGGGELMOPHVIDQAINSLYLGASLRPLI
QPSYSSDMGVGSMTPLHKPPAEGHLSAKDSAAENLLLLAKSASSEKSGSPHS
GQSDYTESNNEKAGVAGSLIYLTNHTITSGVRNGVLPYVKEEQOQYEMRASIEI
ASEGFKVLSGEQVQRAYRCEHCRILFLDHVMTTHMGCHFRDPFECNLCGRHSQDR
YEFSSHMTREHRY"
BASE COUNT 707 a 592 c 597 g 599 t 1 others
ORIGIN
Query Match 33.9%; Score 212.8; DB 4; Length 2496;
Best Local Similarity 65.6%; Pred. No. 5.8e-51;
Matches 341; Conservative 0; Mismatches 152; Indels 27; Gaps 1;
QY 134 gtgcaggccacacagtggaagatgaactgcgagtgtgtgattatctctgcacagt 193
Db 430 GTATCCGCTGCCACAGCGGAAGCAAGTGTGATATCTCTGGGATGATTGCGTGC 489
QY 194 tcaatgtcttaagtctataacgaagccatactggtgaacccattccatgtaatc 253
Db 490 CCAATGTGCTGATGGTGCAACGGAAGCCACACTGGAGAGCGTCCATTCCAGTGC 549
QY 254 agtgggggcattctttactcagaagaagttaacctctccgccacattaaactgcacacag 313
Db 550 AGTGTGCGCTTCTTTCACCCAGAGGCAACCTGCTACCTCACATCAAGCTCCACTCAG 609
QY 314 gggaaaaaaccttttaagtgtcacctctgcacactatgcattgccaaaagaagatgcctca 373
Db 610 GAGAGAAGCCCTTCAAGTGTCACTGTGCAACTATGCTGTCCGCGAGAGACGCCCTCA 669
QY 374 cggggcatcttagacacattctgtggagaaccctacaaatgtgattttgtggaagga 433
Db 670 CGGCTACCTGCGTACCCACTCTGTGGAAACCCCAAGTGTGTACTGTGGGCGGA 729
QY 434 gtacaagcagagaagttcccttgaggagcacaaggagcgtcgcttacattcttcaga 493
Db 730 GCTACAGACGAGCTAGTCTCTTGAGGAGACACAGAGCGGTGTCACTACCTCCAGT 789
QY 494 gcaatgaccacaggggacacatgcgaatgcgagggca----- 529
Db 790 GCATGGGCTCCAGACACAGCATCTATACAGTAGTAAAGGAAGAAAGCAACAGATCAGC 849
QY 529 --agacacataacagcagatgggaagtgaagagctctcgactggacagattagcaa 586
Db 850 AGAGGGAAGACTTAAGCCAGATGGGATCTAAGAGAGCGCTTGGTGCTAGACAGACTAGCTA 909
QY 587 gcaatggcaaaacgaaaaagctcaatgcctcagaatt 626
Db 910 ATAATGTAGCAACAGCTAAGAGCACTATGCCACAGAGTT 949
RESULT 5
CGIKTRF 1557 bp mRNA VRT 03-SEP-1997
LOCUS G.gallus mRNA for Ikaros transcription factor.
DEFINITION
```

```

Qy 386 ggacacattctgtgagaaacccctacaatgtgagttttgtggaaggagttacaaagcaga 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 815 GGACCCACTCCGTTGGTAAACCTCAAAATGGGATATTGTGGCGAAGCTATAAACACAGC 874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 gaagttcccttgagagacacaaggagcgctgcgtacatttcttcagagcactgacc--c 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 875 GAACGCTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGCCCTTC 934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 504 aggggacactgcaagtgcggaggcagacacatacaagca----- 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 935 CGGGCACACTGTACCCAGTCAATTAAGAGAACTAAGCACAGTGAATGCGAGAAGACC 994
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 -----gagatgggaagtgaagagctctcgtactggacagattagcaagcaatgtggca 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 995 TGTGCAAGATAGGATCAGAGAGATCTCGTCTGCTGACAGACTAGCAAGTAATGTGCGCCA 1054
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 599 aacgaaaaagctcaatgcctcagaatt 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1055 AACGTAAGAGCTCTATGCTCAGAAATT 1082
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
HSU40462      HSU40462      3629 bp      mRNA      PRI      30-APR-1996
LOCUS      Human Ikaros/Lyf-1 homolog (hik-1) mRNA, complete cds.
ACCESSION      U40462
NID      91289370
VERSION      U40462.1 GI:1289370
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 3629)
AUTHORS      Nietfeld, W. and Meyerhans, A.
TITLE      Cloning and sequencing of hik-1, a cDNA encoding a human homologue
of mouse Ikaros/Lyf-1
JOURNAL      Immunol. Lett. 49 (1-2), 139-141 (1996)
MEDLINE      96252222
REFERENCE      2 (bases 1 to 3629)
AUTHORS      Nietfeld, W.
TITLE      Direct Submission
Submitted (10-NOV-1995) Wilfried Nietfeld, Department of Virology,
University of Freiburg, Institute for Medical Microbiology and
Hygiene, Hermann-Herder-Strasse 11, Freiburg 79104, Germany
LOCATION/Qualifiers
FEATURES      source
                1..3629
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /tissue_type="bone marrow"
                169..1728
                /gene="hik-1"
                169..1728
                /gene="hik-1"
                /note="similar to mouse Lyf-1, encoded by GenBank
                Accession Number S74708; similar to mouse Ikaros
                DNA-binding protein, Swiss-Prot Accession Number Q03267"
                /codon_start=1
                /product="Ikaros/Lyf-1 homolog"
                /protein_id="AAC50459.1"
                /db_xref="PIR:g1289371"
                /db_xref="GI:1289371"
                /translation="MDADGDMQSVSKESPVPVSDTPDEGDFMPIDPDLSTTSGQ
                OSSKDRVASNVAVETQSDENGACEMGECAEDLRMLDASGERKMGSHRDGSS
                ALSVGGRIPNGKLDICIGIPNVLVHKRSHTERPFCOCNGASFTQKGNL
                LRHLKHSGERPFKCHLCNACRRDALTGLHRTHSVKGPKKCGYCSYKQSSLEE
                HKERKHNLSMGLPGTLYPIVKEETNHEMELDKIGSERSLVLDRLASNVAKRS
                SMPQKFLVDGLSDTPYDSSAYEKENEMKSHYMDQAINNAINYLGAESLRPLVPTP
                PGSEVPVPTSPYQLKHPKLAEGTPRSHSAQSDAVENLLLSKAKLVPSEKASPSN
                SCQDSTSNNEQRGLIYLTNHIAPHARGLSLKEHRAYDLLRAAENSQDALR
                VVSTSGEQMKVYKCEHRCVFLDHWMTIHMGCGRDPECNMGYHSQDRYEFSSH
                ITRGEHRFMS"
BASE COUNT      917 a 1002 c 936 g 773 t 1 others

```

ORIGIN

Query Match 34.2%; Score 214.8; DB 11; Length 3629;
 Best Local Similarity 67.7%; Pred. No. 1.5e-51;
 Matches 344; Conservative 0; Mismatches 137; Indels 27; Gaps 2;

```

Qy 146 ccagtggaagatgaactgcgagtgtgtgtgattatctctgcatacagattcaatgtcttaa 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 506 CTAACGGAAAACTAAAGTGTGATATCTGTGGGATCAITTTGTCATCGGGCCCAATGTGTCTCA 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 206 tgggttaagcagaagcaccatactactggtgaagcccatccagtgtaatacagtggtgggcat 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 566 TGGTTCACAAAAGAAAGCACACTGGAGAACGGCCCTTCCAGTGCATTCAGTGCGGGGCCT 625
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 cttttactcagaaaggtaacctctccgccacattaaactgcacacagggggaaaaacctt 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 626 CATTACCCAGAGAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGGGGGAGAGGCCT 685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 ttaagtgtcactctgcaactatgcatgccaagaagagatgcgtccacgggggcatctta 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 686 TCAAATGCCACTCTGCAACTACGCTGCGCGGAGGAGCGCCCTCAC'TGGCCACCTGA 745
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 ggacacattctgtgagaaacccctacaatgtgagttttgtggaaggagttacaaagcaga 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 746 GGACGCACCTCCGTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAAACAGC 805
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 gaagttcccttgagagacacaaggagcgctgcgtacatttcttcagagcactgacc--c 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 806 GAAGCTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGCCCTTC 865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 504 aggggacactgcaagtgcggaggcagacacacatcaaaagca----- 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 866 CGGSCACACTGTATCCAGTCATTAAAGAAAGAACTAATCACAGTGAATGSCAGAAGACC 925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 -----gagatgggaagtgaagagctctcgtactggacagattagcaagcaatgtggca 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 TGTCAAGATAGGATCAGAGAGATCTCTCGTCTGTCGACAGACTAGCAAGTAATGTGCGCCA 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 599 aacgaaaaagctcaatgcctcagaatt 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 AACGTAAGAGCTCTATGCTCAGAAATT 1013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4

```

LOCUS      OMU92201      2496 bp      mRNA      VPT      22-JAN-1998
DEFINITION      Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, 1x-1 isoform,
                complete cds.
ACCESSION      U92201
NID      92062741
VERSION      U92201.1 GI:2062741
KEYWORDS      rainbow trout.
SOURCE      Oncorhynchus mykiss
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
                Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
                Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE      1 (bases 1 to 2496)
AUTHORS      Hansen, J.D., Strassburger, P. and Du Pasquier, L.
TITLE      Conservation of a master hematopoietic switch gene during
                vertebrate evolution: isolation and characterization of Ikaros from
                teleost and amphibian species
JOURNAL      Eur. J. Immunol. 27 (11), 3049-3058 (1997)
MEDLINE      98056818
REFERENCE      2 (bases 1 to 2496)
AUTHORS      Hansen, J.D.
TITLE      Direct Submission
JOURNAL      Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
                Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
FEATURES      Location/Qualifiers
                source
                1..2496
                /organism="Oncorhynchus mykiss"

```


[illegible]

Search completed: November 6, 1999, 08:05:41
 Job time: 4068 sec


```
Db 290 GGPBARCKCGKVKYNNHFLAIHORSHGTGPRPPKCEGCGFAQKSLQVHTRMHTGERP 349
QY 173 FKCHLCNYA-----
Db 350 YTCVCSKALTTKHSLLEHMSLHSGOKSFCTDCGKYFSONRQLKSHYRVHTGSLPECK 409
2, 182 -CORR-----DALTGHLRTHSVKPYKCEGGRSYKORSSLEEHKEKRCFLQNP----- 231
Db 410 DCHRFMDVSQLKHLRTHGTGKPFCEICGKSFTAKSSLQTHIRHSG--EKPYSCGIC 467
QY 231 --DLGDAASVEARHKAEMGSE-----RALVLDRLASNAVKRKKSSMPQKFIKEKRH 279
Db 468 GKSFSD--SSAKRHCHILHTGKKFCEPCNQLQFARDNLKAHL-----KIHSEKH 517
QY 480 CPDANYPNGMYRKEBEN--MMOTRMMDQAINNAISYLGAFAFRPLVQTP--PAPTSEMP 335
Db 518 ASDASSIGSSNTEEVNRNQLQPOLSTSGEQEIQLLVDSVHNINFMPPGSGISIVT 577
QY 336 VISSVPIALTRADMPGAPQEMEKKRILLPEKILPSE--RGLSPNNSAQDSTDTSNH-- 393
Db 578 AESQNNMTADQANLTLTQOPEQLNLILSAQOEOETHTQSLNMIESQMGPSQTEPVHV 637
QY 393 -----EDROHLYQO-----SHVLPQARNMGMLLKEVPRSFELLKPPP 430
Db 638 ITLSKETLEHLHAHQEOTEELHLATSTSDPAQHLQLTQEPGPPPTHHVP-----OPTP 691
QY 431 I 431
Db 692 L 692

RESULT 11
O43724
ID O43724 PRELIMINARY; PRT; 1207 AA.
AC O43724;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE D29K1.2 (FRAGMENT)
GN D29K1.2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA TUBBY B.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98745; CAB11428.1;
DR PFAM; PF00096; zf-C2H2; 19.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 17.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 1207 AA; 140376 MW; C54700DD CRC32;
```

```
Query Match 11.1%; Score 302.5; DB 4; Length 1207;
Best Local Similarity 21.0%; Pred. No. 3.6e-16;
Matches 111; Conservative 71; Mismatches 207; Indels 139; Gaps 15;

QY 4 IQPTVELKSTEQPLTPESPDALNDYSLPKPHEIENVDSREAPANEDEDGEDSMKVKE 63
Db 744 LQETVLRKKEGSPMSLQSKAQPKYESP-----LESQOQDVETGNEYGNLQKE 794
QY 64 YSDR-----DENIMKPEPMGDAEE-----SEMPYSYARE-----YGDYESIKLER- 104
Db 795 VSEMEPHGTSSRFENDMSKSARCGTREPPEITEPSACSREDKOPTDCNGSVLSEN 854
QY 104 --HVPYDNRPTSGKMNCDVGLSCISFNVLVHVKRSHGTGERPFQCCGASFTQGNLL 161
Db 855 SDHTEHORICPGESYCDGCGAFSGHSLIEHQRIHTGDRPKCECGKAFGRVTLI 914
QY 162 RHILKHTGKPFKHLNACORDALTGHLRTHSVKPYKCEGGRSYKORSSLEEH-- 220
```

```
Db 915 RHKIIHTGKPYKCEGCGAFGRWSALNQHORLHTGKHYHCNDGKAFSQKAGLFHHK 974
QY 220 ---KERCRAFLQNPDLGDAASVEARHKAEMGSE-----RALVLDRLASNAVKRKS 268
Db 975 IHTRDKPYQCTCNKFSRRLTTOHGVHTGAKPYECNECKAFVYN-----SSL 1025
QY 269 MPQKFIGKRRHCFDANYPNGMYKENEMOTRMMDQAINNAISYLGAFAFRPLVOTPPA 328
Db 1026 VSHQBIHHEKCYQC-----KCGKSFQSG-----LIQHORI 1058
QY 329 PTSEMPVISSVPIALTRADMPGAPQEMEKKRILLPEKILPSEGLSPNNSAQDSTD 388
Db 1059 HTGK-PYKCDVCEAF-----IORTSEHQRIHTGERPKYKCDKCGRAFTQR 1105
QY 389 DSNHEDROHLYQOOSHVLPQARNMGMLLKEVPRSFELLKPPPICLRDSTKVINKEGMD 448
Db 1106 SVLTE-----HQRIHT-----GE--R 1119
QY 449 VPRCDCHVFLDYVMFTIHMCGHGFDPFECNMGCGYRSHDRYFESSH 496
Db 1120 PYKDECGNAFRGITSLIQHORIHTGKPYQCEGCGKAFQRQSDLSKH 1167

RESULT 12
O43361
ID O43361 PRELIMINARY; PRT; 691 AA.
AC O43361;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE R30217.1
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA LAMERDIN J.E., MCCREADY P.M., SKORONSKI E., ADAMSON A.W.,
RA BURKHART-SCHULTZ K., GORDON L., KYLE A., RAMIREZ M., STILWAGEN S.,
RA GARNES J., DANGANAN L., POUNDSTONE P., CHRISTENSEN M., GEORGESCU A.,
RA AVILA J., LIU S., BRUCE R., QUAN G., MONTGOMERY M., OW D., NOLAN M.,
RA TRONG S., KOBAYASHI A., OLSEN A.O., CARRANO A.V.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004076; AAB97932.1;
DR PFAM; PF00096; zf-C2H2; 16.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 12.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 691 AA; 80606 MW; 3FEA0BAF CRC32;
```

```
Query Match 11.1%; Score 301.5; DB 4; Length 691;
Best Local Similarity 24.2%; Pred. No. 2.1e-16;
Matches 100; Conservative 60; Mismatches 152; Indels 101; Gaps 16;
```

```
QY 110 SRPTSGKMNCDVGLSCISFNVLVHVKRSHGTGERPFQCCGASFTQGNLRLHKLHTG 169
Db 206 SRPT--PYECTOCGKAFLTQAHVLGVHQKHTGTGEQPYECNCKGKFFMYNSKLIHQKRVHTG 263
QY 170 EKPFGKHLNCAQRRDALTGHLRTHSVKPYKCEGGRSYKORSSLEEHKEKRCRA---F 226
Db 264 ERYECSECGKLFMDSFITLGRHQRVHTGERPFECSCIGKFFSHRSTLNMH--QRVHAGKRL 322
QY 227 LQNPDLGDAASVE---ARHKAEMGS-----ERALVLDRLASNAVKRKKSSMPQKFTG 275
Db 323 YKCEGCGAFSLKHNVOHLKIHTGERPYECCEKAFV-----RKSLVQH--- 370
QY 276 EXRHCFDANYPNGMYKENEMOTRMMDQAINNAISYLGAFAFRPLVOTPPAPSEMP 335
Db 370 QKIHT-DA-----FSKRSDLIQHKRID-----IKRPYTCSECK 403
QY 336 VISSVPIALTRADMPGAPQEMEKKRILLPEKILPSEGLSPNNSAQDSTDTSNHEDR 395
Db 404 AF-----LTQA-----HLVGHQKIHTGER-----PYECTOCAKAFVK 436
```

```

Clogmia albipunctata (mothmidge).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Nematocera; Psychodoidea; Psychodidae; Clogmia.
[1]
RA SEQUENCE FROM N.A.
RP ROHR K.B., TAUTZ D., SANDER K.;
RT "Segmentation gene expression in the mothmidge Clogmia albipunctata
RT (Diptera, Psychodidae) and other primitive dipterans.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ131041; CAA10281.1; -.
KW PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
KW zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 495 AA; 55367 MW; 01F6DECA CRC32;

Query Match 12.6%; Score 342.5; DB 5; Length 485;
Best Local Similarity 23.9%; Pred. No. 6.7e-20;
Matches 127; Conservative 69; Mismatches 171; Indels 165; Gaps 23;

QY 52 DAGESMKVKDYSDRDENIMKP-----EPMGDAESEMPSYAREYSYIKLER 103
DB 37 DSGEDS-----HSPDSDLEPVITDADYDDDAEEDD-----DIRTPKINS 80
QY 104 HVPYDNRPTSGKM-----NCDVGLSCISFNVLVWVKRSH-TGERPFCNOCCASFTQ 156
DB 81 H-----GKMTYKCKODFTAVTKLSF---WEHNRTHIKPEKMLKCKQCFITEY 127
QY 157 KGNLLRHHLKHTGEKPFKCHLNCYACORRDALTGHLRTHSVKEPKYKCFGCRSYKORSLL 216
DB 128 KKHLEYHLRNHNSKPFQCKQCNYSVKNLSHMSHNSNIYQYRCKDCNYATKYCHSL 187
QY 217 EEHKERCRAFLQNDL-----GDAASVEARHKAEMGSEALVLDRLA 259
DB 188 KLRH---LRKYSHPNPVWLYNDGTPNPLRIIDVYGRRGPKVKFKHDEGHH-----NLLN 238
QY 260 SNAVKRKSMPKQFFGEXRHCDFDANYNGYKEKENMMQTMMDQAINNAISYLGAEAF 319
DB 239 SNINTSRSK-----SGKRSDF-----PNF-----EOSQ----- 263
QY 320 RPLVQTPPAPTSEMPVPISSVVPPIALTRADMPGAPQMEKRIILPEK-----ILPSP 374
DB 263 ----HVPTPPSSQALAMPLNLANIQSPSPFLFYLNLFNHILLAQKKAALSQISPSIN 318
QY 375 G-LSPNNSAQDST-----DTSNHDHQHLYQOSHVVYLPQARNG-----MPLL 416
DB 319 GQNEENCNEETPEKEEDPKRMSALDLSNPNSTPTSVQVKH-----KRGKRAFKLELM 373
QY 417 KEVP-----RSFELLKPPPICLR-----DSIKVINKGE-----VMDVF 450
DB 374 KESSDDDEGQTIRTILGEIRS-ELETPKPVQLQLPTSTTTPLKTTSEDDSTSVPELQNL 432
QY 451 RCDHCHVLFLDYVMTIIMGCHGFRDPPECNMCGYRSHDRYEFSSHARGEH 502
DB 433 ECKFCDISFKHVLVITIMHGHYGVNDVFKCNACGKCCEDRVAFFLHIARDAH 484

RESULT 9
O14898 PRELIMINARY; PRT; 506 AA.
AC O14898;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE ZINC FINGER PROTEIN (FRAGMENT).
GN H2F6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98234550.
RA SHANNON M., STUBBS L.;
```

```

"Analysis of homologous XRC1-linked zinc-finger gene families in
human and mouse: evidence for orthologous genes.";
RL Genomics 49:112-121(1998).
DR EMBL; AF027513; AAD12728.1; -.
DR PFAM; PF01352; KRAB; 1.
DR PFAM; PF00096; zf-C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 6.
KW zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 506
SQ SEQUENCE 506 AA; 57498 MW; 7A530D9B CRC32;

Query Match 11.3%; Score 308; DB 4; Length 506;
Best Local Similarity 29.8%; Pred. No. 4.3e-17;
Matches 64; Conservative 33; Mismatches 86; Indels 32; Gaps 3;

QY 29 YSLPKPHEIENVDSREAPANEDEGDSMKVKDEYSDRDENIMKPEPMGDAESEMPSY 88
DB 217 FSCISHHDDNIVHKRDKVHSNSDCGDTLKVSLPTORSHTGQTYQGNCEEA----- 272
QY 89 YAREYSYESIKLHRHVPYDNRPT-----SGKMN--CDVGLS 125
DB 272 ----ENDSSSLHLKHQVHLGKSPACSTHEKDTSSYGIPVQQSVRTGKKRYWCHEGKG 327
QY 126 CISFNVLVWVKRSHITGERPFCNOCCASFTQKGNLLRHHLKHTGEKPFKCHLNCYACORR 185
DB 328 FSQSSNLQTHQRTVTEKPYTCEGKSFNSOSSHLYAHLPIHTGEKPYRCDSCGKGFSSR 387
QY 186 DALTGHLRTHSVKEPKYKCFGCRSYKORSLEEKK 220
DB 388 TDLNIHVRVHTGERPYKCEVCGKGFQSHLQAE 422

RESULT 10
O43167 PRELIMINARY; PRT; 697 AA.
AC O43167;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE KIAA0441.
GN KIAA0441.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA ISHIKAWA K., NAGASE T., NAKAJIMA D., SEKI N., OHIRA M., MIYAJIMA N.,
RA TANAKA A., KOTANI H., NOMURA N., OHARA O.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB007901; BAA23713.1; -.
DR PFAM; PF00651; BTB; 1.
DR PFAM; PF00096; zf-C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 8.
KW zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 697 AA; 78292 MW; 20075420 CRC32;
```

```

Query Match 11.2%; Score 306; DB 4; Length 697;
Best Local Similarity 22.0%; Pred. No. 9.4e-17;
Matches 119; Conservative 69; Mismatches 177; Indels 176; Gaps 18;

QY 2 EDIOTVELKSTEQPLTSPDALNDYSLPKPHEIENVDSREAPANEDEGAD----- 57
DB 217 EESEPTCEPSREEMPVEKD-----ENYD-----PKTEGQASQSRYSKR 256
QY 57 ----SMKVKDEYSDRDENIMKPEPMGDAESEMPSYAREYSYIKLERHVPYDNRSP 112
DB 257 RIWRSVCLKD-----YKLVGQDEHGSAK---RICGRKRKP 289
QY 113 TSGKMNCDVGLSCISFNVLVWVKRSHITGERPFCNOCCASFTQKGNLLRHHLKHTGEK 172
```



```
Db 403 DSTDTESNEDR-----IGVGIPSSGSPQPPTIIVGRHSPAYAKEDPKQEGLRGT 457
QY 430 PICLRDSIKVINKGEVMDVFRCDHCHVFLDYVFTIHMGCHGFRDPFECNMCGYRSHD 489
Db 458 PGPSKEVLRVGSEGPVKAFCHECRILFLDHVFTIHMGCHGFRDPFECNICGYHSOD 517
QY 490 RYEFSSHARGHR 503
Db 518 RYEFSSHIVRGEHK 531

RESULT 4
P79751 PRELIMINARY; PRT; 417 AA.
AC P79751;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE IKAROS-LIKE.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
RN [1]
RP SEQUENCE FROM N.A.
RA GELLNER K., BRENNER S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF056116; AAC34387.1; -.
DR PFAM; PF00096; zf-C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 417 AA; 45664 MW; 7056E12F CRC32;

Query Match 31.8%; Score 866.5; DB 13; Length 417;
Best Local Similarity 37.2%; Pred. No. 3e-62;
Matches 193; Conservative 61; Mismatches 134; Indels 131; Gaps 10;

QY 6 PTVELKSTEEQPLPTESPDALNDYSLPKPHETIENVDSREAPANEDEDAG-----EDSMKV 60
Db 7 PSIOASAEQDFSGGPSATSPNGQOSSPRHSLSVHSIKVELCSDDSPGREGHPRDGGRM 66
QY 61 DEYSDRDENIMKPEMGDAESEMPSYAREYSYDIKLERHVPYDNRPTSCNMCD 120
Db 67 ED--GRRMEDGGRNEVGDDVDGPGROKAGCSSELSPNTASPGPI---RLPNGKLOCE 121
QY 121 VGLSCISFNVLVHKRSHGTGERPQCQCAGASFTQKGNLLRHKLHGTCEKPFKCHLCNY 180
Db 122 ICGIVCTGNVLVHKRSHGTGERPQCQCAGASFTQKGNLLRHKLHGTCEKPFKCHLCNY 181
QY 181 ACORRDALTGHLRTHSVKPYKCEFCGRSYKORSLEEHKERCRAFLQNPD---LGDAAS 237
Db 182 ACRRRDALTGHLRTHAVGKAFKCYSCRSYKQOSTLEDQERCHSYLSRQAAYVVRQAAP 241
QY 238 VEARHIKAEMGSEALVLDRLASVNAKRSKSNPQKFCICEKRCFEDANYPGYMYEKENEM 297
Db 242 GALNMDGMNQNEKTQQVDRLAMTLAKRKTMPQKFLGGS----- 282
QY 298 MOTRMDQAINNAISYLGAERFLVQTPPTAPTSMPVSVSSVPIALTRADMPMGAPQE 357
Db 282 -----TALGREDDQCGCL----- 294
QY 358 MEKKRILLPEKILPSEGLSPNNAQSDTSDTDSNEDR-----QHLYQQSHV 404
Db 294 -----CHAAASP--NGCPDSDTTESTAEQSTRAPAQISNSNHLHYQ--- 335
QY 405 VLPQARNMPLLKVEPRSEFELLKPPICLRDSIKVINKEGEMVDFRCDHCHVFLDYVM 464
Db 335 -----PRG---LHSHSTARETVQVLDREGHVPFRFFHCHLHILFLDHVM 376
QY 465 FTIHMGCHGFRDPFECNMCGYRSHDRYEFSSHIARGHR 503
Db 377 FTIHMGCHGFRDPFECNCGYSSQDRYEFSSHISRGHQ 415
```

```
RESULT 5
O42244 PRELIMINARY; PRT; 338 AA.
AC O42244;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JAN-1999 (Tremblrel. 10, Last annotation update)
DE IKAROS-RELATED TRANSCRIPTION FACTOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RA TURPEN J., KELLEY C., MEAD P., ZON L.;
RL Immunity 0:0-0(1997).
DR EMBL; AF024439; AAB81280.1; -.
DR PFAM; PF00096; zf-C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 2.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
FT NON_TER 328
FT NON_TER 328
SQ SEQUENCE 328 AA; 37116 MW; CCC9E71E CRC32;

Query Match 25.0%; Score 681.5; DB 13; Length 328;
Best Local Similarity 44.4%; Pred. No. 1.9e-47;
Matches 147; Conservative 55; Mismatches 100; Indels 29; Gaps 8;

QY 166 LHTGEKPFKCHLCNYACORRDALTGHLRTHSVKPYKCEFCGRSYKORSLEEHKERCRA 225
Db 1 IRLGEKPFKCHLCNYACORRDALTGHLRTHSVKPYKCEFCGRSYKORSLEEHKERCRA 60
QY 226 FLQNPDLGDAASV-----EARIKAEMGS-----ERALVDRDLASVNAKRSKMP 270
Db 61 YLQNVGMEAGGIVHASPMDCEKNPTMNNMNSLPFERPAVIERLASNMGRKKSSTP 120
QY 271 QKFTIGEK--RHCF--DANYPGYMYEKENEMOTRMDQAINNAISYLGAERFLVQTPP 327
Db 121 QRFGEKLMRYGYDPLHFD--MAYEKEAIIQSQMMDQAINNAITYLQADALRLPIIHS 178
QY 328 APTSEMPVSVSSVPIAL--TRADMPMGA-----POEMEKKRILLPEKILPSEGLSPN 379
Db 179 AAMPEVPPIVSSLYSQVYHPARVERPTSRETSDSNNDMDGPISLRPNHPQEREASPS 238
QY 380 NSAQDSTDTSDNSHEDRQHLVQOQSHVLPQARNMPLLKVEPRSEFELLKPPICLRDSIKV 439
Db 239 NSCLDTTDSSESSHE--APAILPFGNHALNSKRSKQSSVYPKEHGKPLDATKSSLGSKDYVR 297
QY 440 INKEGEMVDFRCDHCHVFLDYVFTIHMG 470
Db 298 FNSEGEQIRAFKCEHCRVFLDHVFTIHMG 328

RESULT 6
O13100 PRELIMINARY; PRT; 138 AA.
AC O13100;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1) (FRAGMENT).
OS IKAROS OR LYF-1.
OS Xenopus laevis/Gillii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98056818.
```


Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2728	100.0	507	11	O08900	O08900 mus musculus
2	1165.5	42.7	537	13	O93581	O93581 brachydactylus
3	1090	40.0	533	11	O92282	O92282 mus musculus
4	866.5	31.8	417	13	P79751	P79751 fugu rubripes
5	681.5	25.0	328	13	O42244	O42244 xenopus laevis
6	546.5	20.0	138	13	O13100	O13100 xenopus laevis
7	345	12.6	594	11	O62514	O62514 mus musculus
8	342.5	12.6	485	5	O96785	O96785 clogmia albipunctata
9	308	11.3	506	4	O14898	O14898 homo sapiens
10	306	11.2	697	4	O43167	O43167 homo sapiens
11	302.5	11.1	1207	4	O43724	O43724 homo sapiens
12	301.5	11.1	691	4	O43361	O43361 homo sapiens
13	298.5	10.9	623	11	O62510	O62510 mus musculus
14	297.5	10.9	604	4	O43309	O43309 homo sapiens
15	297.5	10.9	751	4	O60792	O60792 homo sapiens
16	296.5	10.9	591	11	P70590	P70590 rattus norvegicus
17	296.5	10.9	346	11	O62512	O62512 mus musculus
18	295.5	10.8	650	11	O62886	O62886 rattus norvegicus
19	295	10.8	693	4	O14588	O14588 homo sapiens
20	294.5	10.8	524	11	O88412	O88412 mus musculus
21	294.5	10.8	819	11	O92108	O92108 mus musculus
22	292	10.7	622	4	O15917	O15917 homo sapiens
23	292	10.7	273	4	O92951	O92951 homo sapiens
24	291	10.7	578	4	O15776	O15776 homo sapiens
25	291	10.7	812	11	O35483	O35483 mus musculus
26	289.5	10.6	612	4	O75802	O75802 homo sapiens
27	289	10.6	394	4	O15361	O15361 homo sapiens
28	288	10.6	367	4	O13398	O13398 homo sapiens
29	288	10.6	546	11	O61491	O61491 mus musculus

DR HSP: P08047; LSP2
KW ZINC-FINGER; DNA-BINDING; METAL-BINDING; NUCLEAR PROTEIN.
FT NON_TER 1 1
FT DOMAIN <1 32 KRAB BOX ("A BOX").
FT DOMAIN 33 66 KRAB BOX ("B BOX").
FT DOMAIN 225 457 ZINC-FINGERS.
FT DOMAIN 225 247 C2H2-TYPE.
FT ZN_FING 253 275 C2H2-TYPE.
FT ZN_FING 281 303 C2H2-TYPE.
FT ZN_FING 309 331 C2H2-TYPE.
FT ZN_FING 337 359 C2H2-TYPE.
FT ZN_FING 365 387 C2H2-TYPE.
FT ZN_FING 435 457 C2H2-TYPE.
SQ SEQUENCE 543 AA; 61772 MW; F98854DC CRC32;

Query Match 10.9%; Score 298.5; DB 1; Length 543;
Best Local Similarity 29.9%; Pred. No. 1.5e-13;
Matches 79; Conservative 40; Mismatches 80; Indels 65; Gaps 9;

QY 12 STEQPLPTESP-----DALNDYSLPKPHEIEN-----VDSREA 45
DB 75 SRKEGLPEEPESHVGTREGFPDAPYTLGKDRQCQSOSLALKEQNLLKOLEFGLKEA 134
QY 46 PANEDDAGEDSKVKDEYSDRDENIM---KPEPMGDAESEMPSYAREYSDEY---SI 99
DB 135 PV---QDQAYKTLRLR-----ENCVLSSPNPFPEISRGEYLTYDSQITDSEHNSL 184
QY 100 KLER-----HVPYDMS---RPTSGKM-----NCDVGLSCISFNVLVMVK 136
DB 185 VSQGTSPGQPCGNSDCHRDSSQAIPITELTKSQVDKPKYCTDCGKSFNHAHLTVHK 244
QY 137 RSHGTERPFCNOCGASFTQKGNLLRHKLHTGEKPKCHLCNYACORRDALGTGLRTHS 196
DB 245 RIHTGERPCKEKGAFSONSLVQHERIHTGDKPKYCAECGKSFCHSTHLTVHRRHT 304
QY 197 VEKPYKEFCGGRSYKORSSLEEHK 220
DB 305 GERPYECQDCGRAFNQNSLSLRHK 328

RESULT 15
Z028_XENLA
ID 2028_XENLA STANDARD; PRT; 439 AA.
AC P18747;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DE 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE OOCYTE ZINC FINGER PROTEIN XLCOP28 (FRAGMENT).
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90040698.
RA NIETELD W., EL-BARADI T., MENTZEL H., PIELER T., KOESTER M.,
POETING A., KNOECHEL W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. MOL. BIOL. 208:639-659(1989).
DR PIR; S06556; S06556.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 13.
DR PFAM; PF00096; zf-C2H2; 13.
DR HSP; P08047; LSP2.
KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; REPEAT.
FT NON_TER 1 1
FT ZN_FING 6 28 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
FT ZN_FING 118 140 C2H2-TYPE.
FT ZN_FING 146 168 C2H2-TYPE.
FT ZN_FING 174 196 C2H2-TYPE.
FT ZN_FING 202 224 C2H2-TYPE.

FT ZN_FING 230 252 C2H2-TYPE.
FT ZN_FING 333 355 C2H2-TYPE.
FT ZN_FING 361 383 C2H2-TYPE.
FT ZN_FING 389 411 C2H2-TYPE.
FT ZN_FING 417 439 C2H2-TYPE.
FT NON_TER 439 439
SQ SEQUENCE 439 AA; 51350 MW; 57799C28 CRC32;

Query Match 10.9%; Score 298; DB 1; Length 439;
Best Local Similarity 24.7%; Pred. No. 1.3e-13;
Matches 98; Conservative 53; Mismatches 136; Indels 110; Gaps 15;

QY 96 YESIKLERHVPYDNRSPITSGKMNCDVGLSCISFNVLVMVKRSHSGTGERPFCNOCGASFT 155
DB 129 YPSNLVEHORTHGDRP-----FQTECDKSFIMSKMLVHLRIHTGEKPKYKCECDKSF 184
QY 156 QGNLLRHKLHTGKPKPKCHLCNYACORRDALGTGLRTHSVEKPKYKCEFCGRSYKORSS 215
DB 185 QOSTLVVHORTHGTERPFCSCHEKSFHYAFVHERHTTGEKPKYKCMCDKAYSQRN 244
QY 216 LEEHKERCRAFTLQNPDLGDAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFTG 275
DB 245 LKLLHQ-----KTHESKPKQ----- 259
QY 276 EKRHCDFANYPGYMEKENEMQTRMDQAINNAISYLGAEAFRLVOTPPAPTSEMYVP 335
DB 259 DSPNCEKT-----FEQES-APKATATMDQLHESA-----GLEKVP 291
QY 336 VISSVYPIALTRADMPGAQOEMEKRIILLPEKILPSEKRL-SPNNSAODSTDTSNHHED 394
DB 292 EL-----PEATNSVESP-EAIDDEYKNI--PWSPLSEYLGVLPPPEKQHKCTECCDFLE 344
QY 395 ROHL--YQOSHVVLPQARNGMPLLEKVPSPFELLKPPPICLRDSIKVIN--KEGYMD-- 449
DB 345 KSKLVVHORTH-----GERPEK-----SYCDKTFIRMVHLLHKKIHDGD 386
QY 449 -VFRCDHCHVFLDYVMFTIHMGCHGFRDPFECNMG 484
DB 387 RPYTCAECGKSFIRMSKLTVHRRTHTGERPYTCAECG 423

Search completed: November 6, 1999, 11:46:20
Job time: 6524 sec

```

ZG20_XENLA
ID ZG20_XENLA STANDARD; PRT; 675 AA.
AC P18714;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GASTRULA ZINC FINGER PROTEIN XFG20-1 (XICGF20.1).
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95094815.
RA SCHAEFER U., RAUSCH O., BOWMEISTER T., PIELER T.;
RT "Sequence-specific recognition of a repetitive DNA element by a C2H2
RT zinc-finger protein in Xenopus."
RL EUR. J. BIOCHEM. 226:567-576(1994).
RN [2]
RP SEQUENCE OF 85-613 FROM N.A.
RX MEDLINE; 90040698.
RA NIETZELD W., EL-BARADI T., MENTZEL H., PIELER T., KOESTER M.,
RA POETING A., KNOECHEL W.;
RT "Second-order repeats in Xenopus laevis finger proteins."
RL J. MOL. BIOL. 208:639-659(1989).
RN [3]
RP SEQUENCE OF 85-613 FROM N.A.
RX MEDLINE; 90040698.
RA NIETZELD W., EL-BARADI T., MENTZEL H., PIELER T., KOESTER M.,
RA POETING A., KNOECHEL W.;
RT "Second-order repeats in Xenopus laevis finger proteins."
RL J. MOL. BIOL. 208:639-659(1989).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82643; G576444; -.
DR PIR; S06565; S06565.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 17.
DR PFAM; PF00096; zf-C2H2; 18.
DR HSP; P25490; 12NM.
DR TRANSFAC; T02366; -.
KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; REPEAT.
FT ZN_FING 62 84
FT ZN_FING 90 112
FT ZN_FING 118 140
FT ZN_FING 146 168
FT ZN_FING 174 196
FT ZN_FING 202 224
FT ZN_FING 257 279
FT ZN_FING 286 308
FT ZN_FING 344 366
FT ZN_FING 373 395
FT ZN_FING 424 446
FT ZN_FING 452 474
FT ZN_FING 507 529
FT ZN_FING 535 557
FT ZN_FING 563 585
FT ZN_FING 591 613
FT ZN_FING 619 642
FT ZN_FING 675 711
SQ SEQUENCE 675 AA; 77116 MW; 3468F756 CRC32;

Query Match 11.0%; Score 299.5; DB 1; Length 675;
Best Local Similarity 21.2%; Pred. No. 1.7e-13;
Matches 119; Conservative 88; Mismatches 178; Indels 175; Gaps 24;

QY 14 EEQPLTPESPDALNYSLPK-----PHEIENVDSREAPANEDEDAGDSMKVKDEY 64
DB 5 EEP-----YEPFNTFGNELLPNYQCTGGEALSDTKSDLAYLEVEITDAH 53
OV 65 SDRDENIMKPEMGDAESEMPSYAREYSDESISIKLERHV-PYDNSRPTSGKMNCDVCG 123
DB 54 --EESNTDKPFTC-----TECGKTFTRK-PNYES-----HIRAHKGEKPFs---CMVCD 96
DB 124 LSCISFNVLMVKRSHGTRGPPQCQCAGSFTQKGNLLRHILHTGKPFKCHLCNACQ 183

Db 97 KAFWKSNNLLVHYSVHSGEKPFSCTECDKTFNSKRAQLEKHLRVHTGKPYSCQCGKSF 156
QY 184 RDALTGLRTHSVKPYKCFGSGRSYKORSLSLEHKKRCRAFLQNPDLGDAASVEARHI 243
DB 157 HKCVLDLSHQRTHGDKPFCSTCEGKGFSGRGNLHKHLTKH--LDQP-----HL 203
QY 244 KAEMG---SERALVLDRL-----ASNVAKRSSMPQKFIGEKR--HC 280
DB 204 CAECGKTFESKSTLLEHQIHSIRPLSEFGKTFSDAHLNLLKHQST----FTEQKPFPC 259
QY 281 FDANYNPGMYEKENEMM---QTRMMDQAINNAISYLGAEAFRPLVQTPAPTSMPVPI 337
DB 260 TEC-----GEIFSNHEHLLTHOSTHEQ-----KPFPCTKCWGIF 295
QY 338 SSVPIALTRADMPGAPQEMEKRRILLPEKILPS-ERGLSPNNSAODSTDTDSNHEDRQ 396
DB 296 SNEHELRTHQSTHTGEG-----QKSLPSTESGGTFSEHHEHLLTHQSTHTEQK 342
QY 397 HLYQOQSHVVLPOAR-----NGMPLLEKVPFRSFELLPKPPICLRDSIKVINKEGVM-- 448
DB 343 H-----LPCTEGCGGTFTEQELLAHQSTHTEQKPLP--CTECGIFSDHEHLLTHQ 392
QY 448 -----DVERCDHCHVFLDYVMFTIHGHCCHGFRDPF 478
DB 393 STHTSPSTFVGQTTEDNHQSPSKDHTGKPFSCGKSFYKSVLKDHLVHTGKPY 452
QY 479 ECNMG-----YRSHDR 490
DB 453 HCIEGGRSYTHQSSLSKSHQR 472

RESULT 14
ZIN08_HUMAN STANDARD; PRT; 543 AA.
AC P17098;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 8 (ZINC FINGER PROTEIN HF.18) (FRAGMENT).
GN ZNF8.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATAARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA.
RX MEDLINE; 90169993.
RA LANIA L., DONTI E., PANNUTI A., PASCUCCI A., PENGUE G.,
RA FELICIELLO I., LA MANTIA G., LANFRANCONE L., PELICCI P.-G.;
RT "cDNA isolation, expression analysis, and chromosomal localization of
RT two human zinc finger genes."
RL GENOMICS 6:333-340(1990).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY PRESENT IN MANY HUMAN CELL LINES
CC OF DIFFERENT EMBRYOLOGICAL DERIVATION.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29581; G340448; -.
DR PIR; B34612; B34612.
DR MIN; 194532; -.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 7.
DR PFAM; PF00096; zf-C2H2; 7.
```


KW TRANSCRIPTION REGULATION: NUCLEAR PROTEIN; SPERMATOGENESIS;
KW TRANS-ACTING FACTOR; TESTIS.

FT DOMAIN 1 ASP/GLU-RICH (ACIDIC).

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

FT 78 573 ZINC-FINGERS.

KW Query Match 11.3%; Score 307; DB 1; Length 580;
KW Best Local Similarity 22.5%; Pred. No. 4.2e-14;
KW Matches 139; Conservative 75; Mismatches 204; Indels 200; Gaps 29;

FT 5 OPTVELKSTEQPLTPSPDALNDYSLPKPHEIENVDSREAPANEDEDGDSNKKYDEY 64

FT 26 EPATEEEMTKIGTVTEESGSLB---DVPH-----DSRGKEPFEGLNDQMLFRRQ 77

FT 65 SDRDE-----NIMKP---EPMGDA-----EESEMPY----- 88

FT 78 YNCDECDQSPAWSTGLIRHQRTHWPKYCECGKAFRMSALVLHQRHGTGKPYPCSWC 137

FT 88 --SYARYSDYESIKLER---HVPYDNRPTSGKMNCDVCGLSISFNVLVHVKRSHGT 141

FT 138 IKPSRS--SDL--IKHQRVHTGKPY-----KDCGKAFSOSDLMIHQRIHTG 184

FT 142 ERPFQCNOCGASFTQGNLLRHILHTGKPFKCHLCNYACQRDALTGLRHVSHVEKPY 201

FT 185 EKPYQCHSKSFSQHGVMVHLRIHTGKPYMCMHCYKHFSSOSDLIKHQRHITGKPY 244

FT 202 KCEFCGRSYKQORSLEEHKE-----RC--RAFLQNPDLGDAASVEARHKAEMGS 249

FT 245 KCDVCGKAFSOSDLIRHQRHITGKPYPCAQCNKFSQNSDL-----IKHRRHITGE 297

FT 250 ERALVLDRLASNAVKRSMQPKFEGKRCFDANYNGPYMYKENEMMOTRMMDQAINN 309

FT 298 K-----PYKCEGKAFNQSSVLILHQRHITGKPY-PCNOCTKS-----FSRLSD-LINH 346

FT 310 AISYLGAEARPLVQTPATSEMPVSVISSYPALTRADPM-----GAQOEK-- 361

FT 347 QRHTGK-----PYPCSQCKMFS-----RRSDLVRYHRIHTGKPYEYCDKCG 390

FT 361 -----KRILLPEKILPSEKRLSPNSAOD-----S 385

FT 391 KTFSSQSNLILHQRHITGKPYPCNSCKSFSRGSDLIKHQRVHTGKPYTCNLCSKFS 450

FT 386 TDTDSNHEDROH-----LYQOSHVV-----PQARNGMPLLKEVPRS 422

FT 451 QSSDLTKHQRVHSGKPYHCSCNKAQFQSSDLILHQRVHTGERPYA-----CTQCPRS 504

FT 423 F-----ELLKPPICLRISIKVINKEGEMVDFRCDHCVLFLOYVMTIIMGCHGRDPF 478

FT 505 FSQKSDLIKHQRI-----HTGE--RPHYK--MCGKAFSQCSTLHQRHITGKPY 551

FT 479 ECNMCGYRSHDRYEFSSH 496

FT 552 PCAQCGKFSQRSQSLVNH 569

RESULT 11

233A_HUMAN

ID 233A_HUMAN STANDARD; PRT; 848 AA.

AC Q06730:

DT 01-OCT-1996 (REL. 34, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DE ZINC FINGER PROTEIN 33A (ZINC FINGER PROTEIN K0X31) (K1AA0065)

DE (HA0946) (FRAGMENT).

GN ZNF33A OR ZNF33 OR K0X31.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

RC TISSUE-BONE MARROW;

RX MEDLINE; 96051398.

RA NOMURA N., NAGASE T., MIYAJIMA N., SAZUKA T., TANAKA A., SATO S.,

SEKI N., KAWARABAYASHI Y., ISHIKAWA K.-I., TABATA S.;

"Prediction of the coding sequences of unidentified human genes. II.

RT The coding sequences of 40 new genes (K1AA0041-K1AA0080) deduced by

RT analysis of cDNA clones from human cell line KG-1."

RL DNA RES. 1:223-229(1994).

RN [2]

RP SEQUENCE OF 667-848 FROM N.A.

RX MEDLINE; 93219119.

RA TUNNACLIFFE A., LIU L., MOORE J.K., LEVERSHA M.A., JACKSON M.S.,

FERGUSON-SMITH M.A., THIESEN H.-J., PONDER B.A.;

"Duplicated KOX zinc finger gene clusters flank the centromere of

RT human chromosome 10: evidence for a pericentric inversion during

RT primate evolution."

RL NUCLEIC ACIDS RES. 21:1409-1417(1993).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).

CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-

CC FINGER PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; D31763; G498152; -

CC EMBL; X68687; G829150; -

CC EMBL; X68689; G296455; -

CC PROSITE; PS00028; ZINC_FINGER_C2H2; 16.

CC PFAW; PF00096; zf-C2H2; 16.

CC HSSP; P08045; 12NF.

CC ZINC-FINGER; DNA-BINDING; METAL-BINDING; NUCLEAR PROTEIN; REPEAT.

FT NON_TER 1

FT ZN_FING 366 388

FT ZN_FING 394 416

FT ZN_FING 422 444

FT ZN_FING 450 472

FT ZN_FING 478 500

FT ZN_FING 506 528

FT ZN_FING 534 556

FT ZN_FING 562 584

FT ZN_FING 590 612

FT ZN_FING 618 640

FT ZN_FING 646 668

FT ZN_FING 674 696

FT ZN_FING 702 724

FT ZN_FING 730 752

FT ZN_FING 756 780

FT ZN_FING 786 808

FT ZN_FING 848 AA; 98528 MW; 226A5F84 CRC32;

FT SEQUENCE

SQ

```

RT repressor protein with differential DNA-binding domains." ;
RL DNA CELL BIOL. 14:1971-981(1995).
CC -!- FUNCTION: MAY FUNCTION AS A REPRESSOR OR SILENCER. PROTEIN, AND
CC MOST LIKELY EXERTS ITS REPRESSING ACTIVITY UPON ZINC-DEPENDENT
CC BINDING TO DNA. MAY BE INVOLVED IN PROPER SPERMATOGENESIS BY
CC REPRESENTING THE EXPRESSION OF GENES UNNECESSARY OR INCOMPATIBLE
CC WITH THE MAINTENANCE OF A HAPLOID CELL STATE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: BRAIN, HEART, SPLEEN, THYMUS, AND TESTIS.
CC -!- DEVELOPMENTAL STAGE: THERE IS A MARKED INCREASE AFTER POSTNATAL
CC STAGES 18-20 (SIMULTANEOUSLY TO THE APPEARANCE OF HAPLOID CELL
CC STAGES). MAXIMAL EXPRESSION IS OBSERVED AROUND 2 WEEKS
CC POSTNATALLY, WITH THE EXCEPTION OF BRAIN AND TESTIS. WHERE THE
CC EXPRESSION IS HIGHEST IN EARLIER DEVELOPMENTAL STAGES.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensee@isb-sib.ch).
CC -----
CC EMBL; X79828; G506502; .
CC MGD; MGI:104786; ZFP90.
CC DR PROSITE: PS00028; ZINC_FINGER_C2H2; 13.
CC DR PFAM: PF00096; zf-C2H2; 13.
CC DR HSP: P08047; ISP2.
CC KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN; REPEAT;
CC FT TRANSCRIPTION REGULATION; REPRESSOR.
CC KZ DOMAIN .16 53 KRAB BOX ("A BOX").
CC FT DOMAIN 54 85 KRAB BOX ("B BOX").
CC FT DOMAIN 208 628 ZINC2-FINGERS.
CC FT ZN_FING 208 230 C2H2-TYPE.
CC FT ZN_FING 250 272 C2H2-TYPE.
CC FT ZN_FING 278 300 C2H2-TYPE.
CC FT ZN_FING 306 328 C2H2-TYPE.
CC FT ZN_FING 334 356 C2H2-TYPE.
CC FT ZN_FING 362 384 C2H2-TYPE.
CC FT ZN_FING 390 412 C2H2-TYPE.
CC FT ZN_FING 446 468 C2H2-TYPE.
CC FT ZN_FING 494 516 C2H2-TYPE.
CC FT ZN_FING 522 544 C2H2-TYPE.
CC FT ZN_FING 550 572 C2H2-TYPE.
CC FT ZN_FING 578 600 C2H2-TYPE.
CC FT ZN_FING 606 628 C2H2-TYPE.
CC SEQUENCE 636 AA; 72423 MW; 789G5355 CRC32;
QY Query Match 11.3%; Score 309; DB 1; Length 636;
QY Best Local Similarity 22.9%; Pred. No. 3.5e-14;
QY Matches 116; Conservative 50; Mismatches 140; Indels 200; Gaps 19;
DB QY 49 EDEDAGDSMKVKDEYSDDRD-ENTMKP---EPMGDAEF--SEM-----PY---- 88
DB 153 EQKNVGGSDSLTDLPQLDLSRSSRSDCKTFGNNUHNHNSLVQTQSILAKKKPKCDK 212
QY 88 ---SYAREYSDYESIKLERHPYPDNRSR---PTSGKMN---CDVCGLSCTSFNVLMVHKRSH 139
DB 213 CRKSFIHRSSLINKHEIHKGDPPYNGTDOGAQSRKKHCEACDGKTLFWRTLQLTQHRIH 272
QY 140 TGERPFQNCOGGASTQGNLLRHIIKLHTGEKPKPHLCNTACORRDALTGHLRTHSVKEK 199
DB 273 TGEXPFECNVCGAFRRHSSSLGQHENAHTEGKPKYCSCLGKAFORSSSLVQHQRHTTGEX 332
QY 200 PYKCFEGCRSVKQRSLEEKH-----ERC----- 224
DB 333 PYRCNLGCSRFRHSSTSTQHVETHSGEKPFQCKGKAFCRSCLVSIVQHRTHTGEKPFEC 392
QY 224 ----RAFLQNPDLDGAASVEARHIKAEMGSBRALVLDRLASNVAKRKSSMPOKFTIGEKRH 279

```

```

Db 393 SICRAGQSPSL-----YKHMRI-----HRRKP----- 418
QY 280 CFDANYNPGYMEKENEMQTRMMDQAINNAISYLGAEAFRLVQTPTTAPTSMPVWISS 339
Db 418 -----YQSNFSL-----AFVP--NTP----- 433
QY 340 VYPITALFRAMPGAPOEMEKKRILLPEKILPSPRGSLPNNSAOD-STDTSDNSHEDRQHL 398
Db 433 -----LPQG-----EGLTVEKSYHCNDGKDFGHITDFSEHQRLHA 469
QY 399 YQOSHVVLPQARNCMPLLKVEPRSFELLKPPICLRDSIKVINKEGEVMDVFRCDHCHVL 458
Db 470 GENSY-----GSEQTLLGOOSUHPREKP-----:YQNCVCKA 502
QY 459 FLDYVMFTIHMCGHGRDPFECNMG 484
Db 503 FKRTSFIHHRHTGKPYECNCG 528

RESULT 10
ZF35_MOUSE
ID ZF35_MOUSE STANDARD; PRT; 580 AA.
AC P15620;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 35 (ZFP-35).
GN ZFP35 OR ZFP-35
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MF1; TISSUE-TESTIS;
RX MEDLINE; 90107944.
RA CUNLIFFE V.T., KOOPMAN P., MCLAREN A., TROWSDALE J.;
RT "A mouse zinc finger gene which is transiently expressed during
RL spermatogenesis.";
RL EMBO J. 9:197-205(1990). [2]
RN RP
RP SEQUENCE FROM N.A.
RX MEDLINE; 91065652.
RA CUNLIFFE V., WILLIAMS S., TROWSDALE J.;
RT "Genomic analysis of a mouse zinc finger gene, Zfp-35, that is up-
regulated during spermatogenesis.";
RL GENOMICS 8:331-339(1990).
CC CC
CC -!- FUNCTION: MAY ACT TO CONTROL GENE ACTIVITY DURING THE PACHYTENE
STAGE OF MEIOTIC PROPHASE. MAY FUNCTION AS A TRANSCRIPTION
ACTIVATOR.
CC CC
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC CC
CC -!- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN ADULT TESTIS.
CC CC
CC -!- SIMILARITY: WITH OTHER ZINC-FINGER PROTEINS.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC CC
DR EMBL; X17617; G55473; -
DR EMBL; M36146; G202452; -
DR EMBL; M36145; G202452; JOINED.
DR PIR; A37107; A37107.
DR PIR; S07667; S07667.
DR MGD; MGI:99179; ZFP35.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 17.
DR PFAM; PF00096; ZF-C2H2; 18.
DR HSSP; P25490; 12NM.
DR TRANSFAC; T00921; -
DR KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; DEVELOPMENTAL PROTEIN;

```

QY 140 TGERPFOCNOGASFTOKNLLRLHKLHTEGKPEKCHLNCYACQRRDALTGHLRTHSVK 199
 Db 159 TGERPFOCNOGASFTOKNLLRLHKLHTEGKPEKCHLNCYACQRRDALTGHLRTHSVK 218
 QY 200 PYKCFGRSGRSLKEHKEKRAFLQNPDLGDAASVEARHIAEMGS---ERALVLD 256
 Db 219 PYKCFGRSGRSLKEHKEKRAFLQNPDLGDAASVEARHIAEMGS---ERALVLD 256
 QY 257 RIASNAV---KRKSMPOKF-IGKRHCDFDANYNGYWEKENEMOTRMMDQANNAI 311
 Db 256 REAESSALMKHKKRTHSRPFCSECSRSFTNSDLTAHMKHTEFRNVNLDS----- 310
 QY 312 SYLGAEAFPLVQTPAPTSSEMPVYSSYPALTRADPMGAPQEMEKRLILLPEKILP 371
 Db 310 SYLGAEAFPLVQTPAPTSSEMPVYSSYPALTRADPMGAPQEMEKRLILLPEKILP 371
 QY 372 SERGLSPNNSAQDSTDSDNHE---DROHLYQOSHVLVPOARNGMPLLKEVPRSFELL-- 427
 Db 343 SERGLSPNNSAQDSTDSDNHE---DROHLYQOSHVLVPOARNGMPLLKEVPRSFELL-- 427
 QY 427 KPPLICLRDSIKVINEGEMDV-----FRCDHCVLFVDYVNFTHMGCHGFRDPF 478
 Db 379 ERPYQCAECHKGFIQKSLVHLRTHTEGKPEKCHLNCYACQRRDALTGHLRTHSVK 438
 QY 479 EGNMCGYRSHDRYFSSH--IARGE 501
 Db 439 KSCDCGKEFTQSNLILHORIHTGE 463

RESULT 8
 HUNB_TRICA STANDARD; PRT; 524 AA.
 AC Q01791;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE HUNCHBACK PROTEIN.
 GN HB.
 OS TRIBOLIUM CASTANEUM (RED FLOUR BEETLE).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; COLEOPTERA; POLYPHAGA; CUCUJIFORMIA; TENEBRIONIDAE;
 OC TRIBOLIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96125163.
 RA WOLFF C., SOMMER R., SCHRODER R., GLASER G., TAUTZ D.;
 RT "Conserved and divergent expression aspects of the Drosophila
 RT segmentation gene hunchback in the short germ band embryo of the
 RT flour beetle Tribolium".
 RL DEVELOPMENT 121:4227-4236(1995).
 RN [2]
 RP SEQUENCE OF 243-311 FROM N.A.
 RX MEDLINE; 93066327.
 RA SOMMER R.J., RETZLAFF M., GOERLICH K., SANDER K., TAUTZ D.;
 RT "Evolutionary conservation pattern of zinc-finger domains of
 RT Drosophila segmentation genes".
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:10782-10786(1992).
 CC -!- FUNCTION: HUNCHBACK IS A GAP CLASS SEGMENTATION PROTEIN.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X91618; G1008447; -.
 DR EMBL; L01615; G16162; -.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
 DR PFAM; PF00096; zf-C2H2; 6.
 DR HSP; P08151; 2GLI.

KW DEVELOPMENTAL PROTEIN; GAP; ZINC-FINGER; METAL-BINDING; DNA-BINDING;
 KW REPEAT; NUCLEAR PROTEIN.
 FT DOMAIN 202 311 ZINC-FINGERS I.
 FT DOMAIN 471 523 ZINC-FINGERS II.
 FT ZN_FING 202 224 C2H2-TYPE.
 FT ZN_FING 231 253 C2H2-TYPE.
 FT ZN_FING 259 281 C2H2-TYPE.
 FT ZN_FING 298 311 C2H2-TYPE.
 FT ZN_FING 471 493 C2H2-TYPE.
 FT ZN_FING 499 523 C2H2-TYPE.
 SQ SEQUENCE 524 AA; 59514 MW; 8B8F65BD CRC32;

Query Match 11.4%; Score 310; DB 1; Length 524;
 Best Local Similarity 23.5%; Pred. No. 2.3e-14;
 Matches 124; Conservative 66; Mismatches 192; Indels 146; Gaps 20;

QY 5 OPTVELKSTEQP-----LPTESPDALNDYSLPKPHEIENVDSREAPANEDEAGEDSM 58
 Db 112 EPLVSPKSEKEEMETTLTPCASPKNRPD-----DNQDLRLRLMSLEKSGLESS 162
 QY 59 KVK-----DEYSDRDENIMKPEPMGDAEEMSPYSAREYSDYSEIKLERHVPYDNRPTS 114
 Db 163 KTSEHSDVDELSGSDN-----DAEE-----YDEQSLR-----VPKYNHSGKI 199
 QY 115 GKMDVCGGLSCIS-----FNVLMVHKRSHTGERPFOCNOGASFTOKNLLRLHKLHTEG 170
 Db 200 KTFKCKQCFVAITKLEQNHSHKVHIRE---DKRLTCTPKCFITEYKHHLEYHLRNHAGS 256
 QY 171 KPFKCHLNCYACQRRDALTGHLRTHSVKPYKCFGRSGRSLKEHKEKRAFLQNP 230
 Db 257 KPFOCNKCDYTCVKNKMLNSHMSHNVYRSCDCSVATKYCHSLKILHRR----- 309
 QY 231 DLGDAASVEARHIAEMGSERALVLDLRLASNVAKRSMPOKFIGEKHRCFDANYPGM 290
 Db 309 -----YCHTENVVLDE-----EGNPCPDII 328
 QY 291 YE-----KENEMQTR-MMDQAINNAISYLGAEAFRPLVQTP---PAPTSSEMPVYSSVYP 342
 Db 329 IDVHGTRGRPKIKTPKAAEAKPETLPFLNLQQLPFGYPPFGFPPNAQLLQQL----- 384
 QY 343 IALTRADPMGAPQEMEKRLILLPEKILPSEGLSPNNS-----AQDSTDSDNHEDRQ- 397
 Db 384 ---IRERLAVGSGQ---EESRVLDSLKPGCVTGEOKSRKGPAPKVDPTQVESEDEET 439
 QY 397 ---HLYQQSHVVLVPOARNGMPLLKEVPRSFELLKPPICLRDSIKVINKEGEVMDVFRCDH 454
 Db 440 STTVFSNVVQEEAK-----KEE-----SDSNNNNKE-----EGNSCQY 475
 QY 455 CHVLFLDYVMTIIMGCHGFRDPFCNMGYRSHDRYFSSHIAARGEH 502
 Db 476 CNIAFGDAVLYTIHMGYHGHNPFTCNMCGVECDKVSFFLHIARVSH 523

RESULT 9
 ZF90_MOUSE STANDARD; PRT; 636 AA.
 AC Q61967;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE ZINC FINGER PROTEIN 90 (ZF90-90) (ZINC FINGER PROTEIN NK10).
 OS MUS MUSCULUS (MOUSE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6; TISSUE-BRAIN;
 RX MEDLINE; 96069544.
 RA LANGE R., CHRISTOPH A., THIESEN H.-J., VOPPER G., JOHNSON K.R.,
 RA LEMAJRE L., PLOMANN M., CREMER H., BARTHELS D., HEINLEIN U.A.O.;
 RT "Developmentally regulated mouse gene NK10 encodes a zinc finger

RA HAHM K., COBB B.S., MCCARTY A.S., BROWN K.E., KLUG C.A., LEE R.,
RA AKASHI K., WEISSMAN I.L., FISHER A.G., SMALE S.T.;
KI "Helios, a T cell-restricted Ikaros family member that quantitatively
RT associates with Ikaros at centromeric heterochromatin";
RL GENES DEV. 12:782-796(1998).
CC -1- FUNCTION: ASSOCIATES WITH IKAROS AT CENTROMERIC HETEROCHROMATIN.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE T-CELL LINEAGE. ABUNDANT IN
CC THYMUS, LOW EXPRESSION IN BONE MARROW AND BRAIN AND
CC NO DETECTABLE EXPRESSION IN SPLEEN, LIVER, KIDNEY OR MUSCLE.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN; HELIOS A AND
CC B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF044257; G2829277; -
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
CC DR HSP; P08047; 1SP2.
CC DR
CC KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
CC FT DOMAIN 112 219 ZINC-FINGERS I.
CC FT DOMAIN 471 523 ZINC-FINGERS II.
CC FT ZN_FING 112 134 C2H2-TYPE.
CC FT ZN_FING 140 162 C2H2-TYPE.
CC FT ZN_FING 168 190 C2H2-TYPE.
CC FT ZN_FING 196 219 C2H2-TYPE.
CC FT ZN_FING 471 493 C2H2-TYPE.
CC FT ZN_FING 499 523 C2H2-TYPE.
CC FT VARSPLIC 111 136 MISSING (IN HELIOS A).
CC FT SEQUENCE 526 AA; 59388 MW; AD142BD4 CRC32;
CC
CC
CC Query Match 45.2%; Score 1234; DB 1; Length 526;
CC Best Local Similarity 49.7%; Pred. No. 6.7e-78;
CC Matches 248; Conservative 67; Mismatches 138; Indels 46; Gaps 8;

QY 34 PHEINVDSEAPANEDEDAGESMKVKDEYSDRDENIMKPEPMGDAESEMYSYAREY 93
DB 41 PSHMTSTNSVKLEMQSDECDRPLSREDEIRGHDEGSLEEPL--IESSEVADN--RKV 96
QY 94 SDYESIKLERHVPYDNRPTSGKMNCDVGLCTISFNVLAVHRSHTGERPFCNOCGAS 153
DB 97 QDLOG-----EGGRLPNGKLKCDVCGMVCIGPNVLVHRSHTGERPFCNOCGAS 148
QY 154 FTQGNLLRHLKHTGKPKFCHLCNACORRDALTLGHLRTHSVKPKCFGCGRSYKOR 213
DB 149 FTQGNLLRHLKHSKPKFPCPCSVACRRDALTGLRTHSVGPKHCNCGRSYKOR 208
QY 214 SSLSEHKERCAFQNLDPDGLDAASVEARHTK-----AEMGSEALVLD 257
DB 209 SSLSEHKERCHNYLQNVSMGAAGVMSSHVPPMEDCKEQEPIIMDNNTLSLVPFFERPAVIEK 268
QY 258 LASNVAKRKSMQPKTGERKHF--DANYNFCYMEKENEMKMTQMDQAINNATSYL 314
DB 269 LTANMGARKSTPOKFGVKLMRFSPDIHFDNMLTKEALQSHMDQAINNATSYL 328
QY 315 GAERFLVQTPPTAPTEMVPISSVY-----PIALTRADMPGAPOMEKKRI 363
DB 329 GAELHPLMHAPSTIAEVAPVISSAYSQVYHPNRIERPISRETSD--SHENNMDGPTS 365
QY 364 LLPKILPISERGLSPNNSAODSTDSDNHDRLHLYQOSSHVLPQANGPMLKKEVPRSF 423
DB 386 LIRPKSRPQERASPSNCSLDSDSDSDHDDROS-YQGNPALNPKRKQSPAYMKEDVKAL 444
QY 424 ELLKPPPICLRDLSKVINKEGVMDFVRCDBCHVLFVDYVMTIHMSCHGFDPFECNMC 483

DB 445 DATKAPKSLKDIYKVFNGEGEQIRAFKCEHCRVLFLOHVMYTIHMCHGYRDPLECNIC 504
QY 484 GYRSHDRYEFSSHARGEH 502
DB 505 GYRSQDRYEFSSHIVRGH 523
RESULT 5
IKAR_ONCMY STANDARD; PRY: 522 AA.
AC OL3089; OL3088; OL3086; OL3087;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS.
GN IKAROS.
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIARDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;
OC SALMONIDAE; ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SHASTA; TISSUE-THYMOCYTES;
RX MEDLINE; 98056818.
RA HANSEN J.D., STRASSBURGER P., DU PASQUIER L.;
RT "Conservation of a master hematopoietic switch gene during vertebrate
RT evolution: isolation and characterization of Ikaros from teleost and
RT amphibian species.";
RL EUR. J. IMMUNOL. 27:3049-3058(1997).
CC -1- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES
CC EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: EIGHT FORMS OF THE PROTEIN; IK1-IK8; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT
CC OF IK1.
CC -1- TISSUE SPECIFICITY: EXPRESSION MAINLY LIMITED TO THYMUS, SPLEEN,
CC AND PRONEPHROS. VERY LOW EXPRESSION IN LIVER. NO EXPRESSION IN
CC TESTIS, BRAIN, EYE AND MUSCLE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AT DAY 3-4 IN THE YOLK SAC
CC AND AT DAY 5-6 IN THE EMBRYO PROPER.
CC -1- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U92201; G2062742; -
CC DR EMBL; U92200; G2062740; -
CC DR EMBL; U92198; G2062736; -
CC DR EMBL; U92199; G2062738; -
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
CC DR PFAM; PF00096; zf-C2H2; 5.
CC KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
CC FT DOMAIN 125 203 ZINC-FINGERS I.
CC FT DOMAIN 468 520 ZINC-FINGERS II.
CC FT ZN_FING 125 147 C2H2-TYPE.
CC FT ZN_FING 153 175 C2H2-TYPE.
CC FT ZN_FING 181 203 C2H2-TYPE.
CC FT ZN_FING 209 232 C2H2-TYPE.
CC FT ZN_FING 468 490 C2H2-TYPE.
CC FT ZN_FING 496 520 C2H2-TYPE.
CC FT VARSPLIC 55 148 MISSING (IN IK-2, IK-4 AND IK-8).
CC FT VARSPLIC 55 291 MISSING (IN IK-6).
CC FT VARSPLIC 149 291 MISSING (IN IK-5).

Db 344 ISPMQL-----HKPLGDNQTSNHTAODSAVENLLLSKAKSVSSERDASFSNSQDST 398

Qy 387 DTDSDHEDRQHYQOQSHVVLPOARNGMPLLEKVPFSFELLKPPPTCLRDSIKVINKEGV 446

Db 399 DTSNEERSGLIYLTNIGHPHARGIS-VKEESROFDVLRAGTDNSQDAFKVSSNCEQ 457

Qy 447 MDVFRCDHCHVLFDYVMTTHMGCHGFRDPFECNMGYSHDRYEFSSHARGHR 503

Db 458 VRYVYKCEHCRVFLDHVMTTHMGCHGFRDPFECNMGYSHDRYEFSSHARGHR 514

RESULT 3

IKAR_MOUSE

ID IKAR_MOUSE STANDARD; PRT; 517 AA.

AC Q03267; Q64044; Q64045; Q64051;

DT 01-OCT-1993 (REL. 27, CREATED)

DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1).

GN IKAROS OR LYF1.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 5).

RC TISSUE-EMBRYO;

RX MEDLINE; 93068267.

RA GEORGOPOULOS K., MOORE D.D., DERFLER B.;

RT "Ikaros, an early lymphoid-specific transcription factor and a

RT putative mediator for T cell commitment.";

RL SCIENCE 258:808-812(1992).

RN [2]

RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.

RX MEDLINE; 95021239.

RA HAHM K., ERNST P., LO K., KIM G.S., TURCK C., SMALE S.T.;

RT "The lymphoid transcription factor Lyf-1 is encoded by specific,

RT alternatively spliced mRNAs derived from the ikaros gene.";

RL MOL. CELL. BIOL. 14:7111-7123(1994).

CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF

CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE

CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL

CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)

CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED

CC DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- ALTERNATIVE PRODUCTS: SIX FORMS OF THE PROTEIN, ISOFORMS I-VI, ARE

CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT

CC OF ISOFORM VI.

CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN T-CELLS AND THEIR

CC PROGENITORS, AND ALSO IN B-CELLS.

CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L03547; G198287; -

DR EMBL; S74517; G807149; ALT_SEQ.

DR EMBL; S74518; G807151; -

DR EMBL; S74708; G807153; -

DR MGD; MGI:96535; IKAROS.

DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.

DR PFAM; PF00096; zf-C2H2; 4.

DR HSP; P08047; ISP2.

DR TRANSFAC; T01470; -

KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;

KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.

FT DOMAIN 117 223 ZINC-FINGERS I.

FT DOMAIN 457 479 ZINC-FINGERS II.

FT ZN_FING 117 139 C2H2-TYPE.

FT ZN_FING 144 166 C2H2-TYPE.

FT ZN_FING 172 194 C2H2-TYPE.

FT ZN_FING 200 223 C2H2-TYPE.

FT ZN_FING 457 479 C2H2-TYPE.

FT VARSPLIC 53 53 M -> VAYGADGRDFPHALISDRGM (IN ISOFORMS

FT VARSPLIC 54 140 II AND IV).

FT VARSPLIC 54 140 MISSING (IN ISOFORM V).

FT VARSPLIC 54 282 MISSING (IN ISOFORMS I AND II).

FT VARSPLIC 141 282 MISSING (IN ISOFORMS III AND IV).

FT CONFLICT 234 235 VC -> MY (IN REF. 2).

FT CONFLICT 480 482 MISSING (IN REF. 2).

Qy 12 STEOPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGEDSMKVDEYSDRDENI 71

Db 13 SGKESPPVSDTPDE-GDEPMPVPEDLSTTSGAQQNSKSDRGMSN---VKVETQSDENG 68

Qy 72 MKPEPMG-----DAESEMPPYAYREYSDYESIKLERHVPYDNSRPTSGKMCNDV 121

Db 69 RACEMNGEECAEDLRMLDASGERKMGSHRQGGSSALS-----GVGGIRLPNGKLCDCI 121

Qy 122 CGLSCISFNVLWVHKRSHTGERPFOCNOCCASFTQKGNLLRHKLHTGERPKFCHLCNYA 181

Db 122 CGVICGPNVLWVHKRSHT-ERPFOCNOCCASFTQKGNLLRHKLHTGERPKFCHLCNYA 180

Qy 182 CORRDAITGHLRTHSVKPKYKCEFCGRSYKQRSLEEHKRCRAFLQNPDLGDAASY--- 239

Db 181 CRRRDALTGHLRTHSVGPKHGKGYGRSYKQRSLEEHKRCRHNLYESMGLPGVCPVIKE 240

Qy 239 EARHIK-----AEMGSRALVLDRLASNAVKRSMPOKFIGEKRRHCF-DANYNGPYME 292

Db 241 ETNENMAEDLCKTAGERSLVLDRLASNAVKRSMPOKFIGLGDK--CLSDNPDYDSA-NYE 297

Qy 293 KENEMOTRMDOAINNAISVLGAEARPLVOTPPPTSEMPVVISVYPITALTRAMP- 352

Db 298 KE-DWMTSHVQDQAINNAIYLGAEURLVQTPPG-SSEVPVVISVYSHKPPSGOPP 355

Qy 352 ---MGAPQEMEKRLPEKILPSEGLSPNNSAQDSTDSDSNHEDRQH--LYQQSHVYL 406

Db 356 RSNHSAQDAVDNLLLSKAKSVSSERASPSNSQDSTDSTESNAEQSGLIYLTNH-IN 414

Qy 407 POARNGPLLEKVPFSFELLKPPPTCLRDSIKVINKEGVMDVPRCDHCHVLFDYVMT 466

Db 415 PHARNGL-ALKEQRAYEVLRAASENSQDAFRVVVSTSGEQLKVYKCEHCRVFLDHVMT 473

Qy 467 IHM---GCHFRDPFECNMGYSHDRYEFSSHARGHR 503

Db 474 IHMGCHGCHFRDPFECNMGYSHDRYEFSSHARGHR 513

RESULT 4

HELI_MOUSE

ID HELI_MOUSE STANDARD; PRT; 526 AA.

AC P81183;

DT 13-JUL-1998 (REL. 36, CREATED)

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE ZINC FINGER PROTEIN HELIOS.

GN HELIOS.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 246-267 AND 289-306.

RC STRAIN-BALB/C; TISSUE-THYMUS;

RX MEDLINE; 98180981.

```

FT VARSPLIC 10 53 MISSING (IN IK4).
FT VARSPLIC 54 140 MISSING (IN IK2).
FT VARSPLIC 54 283 MISSING (IN IK6).
FT VARSPLIC 197 283 MISSING (IN IK3 AND IK4).
FT VARSPLIC 141 283 MISSING (IN IK5).
FT CONFLICT 11 12 QV -> FS (IN REF. 2).
FT CONFLICT 214 214 S -> T (IN REF. 2).
FT CONFLICT 245 245 N -> K (IN REF. 2).
FT CONFLICT 296 296 MISSING (IN REF. 2).
FT CONFLICT 298 298 S -> T (IN REF. 2).
FT CONFLICT 352 355 KPLA -> RRS (IN REF. 2).
FT CONFLICT 372 372 N -> Y (IN REF. 2).
FT CONFLICT 420 426 PHARGL -> RRAQV (IN REF. 2).
SQ SEQUENCE 519 AA; 57528 MW; 5A97272D CRC32;

Query Match 49.1%; Score 1339.5; DB 1; Length 519;
Best Local Similarity 53.6%; Pred. No. 3.6e-85;
Matches 279; Conservative 71; Mismatches 124; Indels 47; Gaps 14;

QY 12 STEEQPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGESMKYKDEYSDRDENI 71
Db 13 SGKESPPVSDTDE-GDEPMPIPEDISTTSGQGSKSDRVASN---VKVETQSDENG 68

QY 72 MKPEPMG-----DAESEMPPSYAREYSDYSEIKLERHVPYDNRPTSGKMNCDV 121
Db 69 RACEMNGECAEDLMDLDSGKMGSHRDQGSALU-----GVGGIRLPNGKLUKCDI 121

QY 122 CGLUSCIFNVLMVHKRSHTGERFQCNQCGASFQKGNLLRHLIKHTGKPKFCHLCNTYA 181
Db 122 CGIICGPNVLMVHKRSHTGERFQCNQCGASFQKGNLLRHLIKHSGEKPKFCHLCNTYA 181

QY 182 CQRDALTGHLRTHSVKPKYKCFECGRSYKQRSLSLEHKERCRAFLQNPDL-GDAASV-- 239
Db 182 CRRDALTGHLRTHSVKPKYKCFECGRSYKQRSLSLEHKERCRAFLQNPDL-GDAASV-- 239

QY 239 -EARHIK-----AEMGSEALVLDRLASNVAKRKSSMPKOFICEKRCRFDANNPGMYE 292
Db 242 EETNHSEMAEDLCKIGERSLVLDRLASNVAKRKSSMPKOFICEKRCRFDANNPGMYE 292

QY 293 KENEMOTRMDOAINNAYSILGAERPLVQTPPTSEMPVSVSYPIALTRADMPM 352
Db 301 KENEMKSHVMDQAINNAYSILGAERPLVQTPPTG-GSEVPVSPMYQLHKPLAE--- 357

QY 353 GAPOE-----MEKRILLPEKILPSERGLSPNNQAQSDTSDTSDNHDQRH--LYQOS 402
Db 357 GTPRNSHSAQDSAVENLLLSKAKLPSEAREASPSNQCQSDTSDTESNNEEQRSGLIYLTN 416

QY 403 HVLVPOARNGMPLKKEVPRSFELLPKPPICLRDSIKVINKEGEVMDVFCDCRCHVFLDY 462
Db 417 HIA-PHARGLS-LKEHRAVDLLRASENSQDALRVSTSGQMVKYKCEHRCRVFLDH 474

QY 463 VMTIHWGCHGFRDPPECNMGYRSHDRYEFSSHARGHR 503
Db 475 VMTIHWGCHGFRDPPECNMGYRSHDRYEFSSHARGHR 515

RESULT 2
IKAR_CHICK ID IKAR_CHICK STANDARD; PRT; 518 AA.
AC 042410;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS.
GN IKAROS OR IK.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H.B2; TISSUE=THYMUS;
RX MEDLINE; 97439462.

```

```

RA LIIPPO J., LASSILA O.;
RT "Avian Ikaros gene is expressed early in embryogenesis.";
RL EUR. J. IMMUNOL. 27:1853-1857(1997).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES
CC EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC HEMATOPOIETIC ORGANS
CC SUCH AS THE BURSA OF FABRICIUS, THYMUS AND SPLEEN. IN THE ADULT,
CC EXPRESSED IN SPLEEN, THYMUS, BURSA AND PERIPHERAL BLOOD
CC LEUKOCYTES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYO FROM EMBRYONIC DAY
CC 2 ONWARDS.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y11833; E309211;
CC DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
CC DR PFAM: PF00096; zf-C2H2; 5.
CC KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
CC FT DOMAIN 117 224 ZINC-FINGERS I.
CC FT DOMAIN 461 513 ZINC-FINGERS II.
CC FT ZN_FING 117 139 C2H2-TYPE.
CC FT ZN_FING 145 167 C2H2-TYPE.
CC FT ZN_FING 173 195 C2H2-TYPE.
CC FT ZN_FING 201 224 C2H2-TYPE.
CC FT ZN_FING 461 483 C2H2-TYPE.
CC FT ZN_FING 489 513 C2H2-TYPE.
CC SQ SEQUENCE 518 AA; 57586 MW; ABADF123 CRC32;

Query Match 48.4%; Score 1321.5; DB 1; Length 518;
Best Local Similarity 52.0%; Pred. No. 6.3e-84;
Matches 279; Conservative 66; Mismatches 131; Indels 61; Gaps 13;

QY 2 EDIQTVELKSTEEQPLTESPDALNDYSLPKPHEIE-NVDSREAPANEDEDAGESMKV 60
Db 4 DEAQDMSQVSGKSPISDVDPDA--DEPMVPEDLSTTTGGQSQVKNERVLAVG----NI 57

QY 61 KDEYSDRDENIMKPEPMG-----DAESEMPPSY-----AREYSDYSEIKLERHVP 106
Db 58 KIETOSDEENGACEMNGEGEACEDLMDLDSGKMGSHNGPGSKAMSGVGIRL----- 113

QY 107 YDNRPTSGKMNCDYCGLSICISFNVLVHKRSHTGERPFQCNQCGASFQKGNLLRHLIKL 166
Db 113 -----PNGKLCDCIGIICGPNVLMVHNRSHTGERPFQCNQCGASFQKGNLLRHLIKL 166

QY 167 HTGKPKFCHLCNYACORRDALTGHLRTHSVKPKYKCFECGRSYKQRSLSLEHKERCRAF 226
Db 167 HSGKPKFCHLCNYACRRRDALTGHLRTHSVGKPKYKCFECGRSYKQRSLSLEHKERCCHNY 226

QY 227 LQNPDLG-----DAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPKOFICEK 277
Db 227 LQTMISNNLSYVKEETNQSMEDLCKIGERSLVLDRLASNVAKRKSSMPKOFICEK 286

QY 278 RHCF-DANNPGMYEKENEMOTRMDOAINNAYSILGAERPLVQTPPTSEMPV 336
Db 287 --CLSDLPYDATNTYKENEIMQTHVIDQAINNAYSILGAERPLVQTPPV-GSEVPV 343

QY 337 ISSVPIALTRADMPMGAPO-----EMEKRILLPEKILPSERGLSPNNQAQSDT 386

```

C:\protein - protein search, using sw model
 Run on: November 6, 1999, 09:57:36 ; Search time 13.55 Seconds
 (without alignments)
 1057.714 Million cell updates/sec
 GenCore version 4.5
 Copyright (c) 1993 - 1998 Compugen Ltd.
 Title: US-09-019-348-2
 Perfect score: 2728
 Sequence: 1 MEDIQTVELKSTEEQPLPT.....HDIYFSSHTARGEHRAMLK 507
 Scoring table: BLOSUM62
 Searched: 77977 seqs, 28268293 residues
 Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339.5	49.1	519	1 IKAR_HUMAN	Q13422 homo sapien
2	1321.5	48.4	518	1 IKAR_CHICK	Q42410 gallus gall
3	1279.5	46.9	517	1 IKAR_MOUSE	Q03267 mus musculus
4	1234	45.2	526	1 HELI_MOUSE	P81183 mus musculus
5	1211	44.4	522	1 IKAR_ONCMY	Q13089 oncorhynchu
6	344	12.6	553	1 2F37_MOUSE	P17141 mus musculus
7	316	11.6	1350	1 XF1N_XENLA	P08045 xenopus lae
8	310	11.4	524	1 HUNB_TRICA	Q01791 tribolium c
9	309	11.3	636	1 2F90_MOUSE	Q61967 mus musculus
10	307	11.3	580	1 2F35_MOUSE	P15620 mus musculus
11	302.5	11.1	848	1 233A_HUMAN	Q06730 homo sapien
12	302.5	11.1	647	1 2G48_XENLA	P18723 xenopus lae
13	299.5	11.0	675	1 2G20_XENLA	P18714 xenopus lae
14	298.5	10.9	543	1 2N08_HUMAN	P17098 homo sapien
15	298	10.9	439	1 2028_XENLA	P18747 xenopus lae
16	287.5	10.9	595	1 2N85_HUMAN	Q03923 homo sapien
17	286.5	10.9	510	1 CF2_DROME	P20385 drosophila
18	296	10.9	407	1 OZF_RAT	O62981 rattus norv
19	295.5	10.8	803	1 2N43_HUMAN	P28160 homo sapien
20	293.5	10.8	393	1 MLZ4_MOUSE	Q03309 mus musculus
21	293.5	10.8	614	1 2F29_MOUSE	Q07230 mus musculus
22	293.5	10.8	485	1 2N42_HUMAN	P28698 homo sapien
23	292.5	10.7	321	1 2177_HUMAN	Q13360 homo sapien
24	292.5	10.7	1191	1 2N91_HUMAN	Q05481 homo sapien
25	292	10.7	347	1 KR2_MOUSE	P08043 mus musculus
26	292	10.7	407	1 OZF_MOUSE	Q62513 mus musculus
27	291	10.7	379	1 211B_HUMAN	Q06732 homo sapien
28	290.5	10.6	726	1 2184_HUMAN	Q99676 homo sapien
29	289.5	10.6	626	1 2189_HUMAN	Q07580 homo sapien
30	288.5	10.5	519	1 2N35_HUMAN	P13682 homo sapien
31	286.5	10.5	292	1 OZF_BOVIN	Q28151 bos taurus
32	286.5	10.5	292	1 OZF_HUMAN	Q15072 homo sapien
33	286.5	10.5	738	1 2N84_HUMAN	P51523 homo sapien
34	286	10.5	572	1 KID1_MOUSE	Q61751 mus musculus
35	286	10.5	629	1 2195_HUMAN	O14628 homo sapien
36	286	10.5	428	1 2F26_MOUSE	P10076 mus musculus
37	284.5	10.4	506	1 2157_HUMAN	P51786 homo sapien
38	282	10.3	474	1 2141_HUMAN	Q15928 homo sapien
39	281	10.3	247	1 2020_XENLA	P18744 xenopus lae
40	280.5	10.3	682	1 2N45_HUMAN	Q02386 homo sapien
41	278	10.2	794	1 2151_MOUSE	O60821 mus musculus
42	277.5	10.2	542	1 2N41_HUMAN	P51814 homo sapien
43	277	10.2	469	1 2135_HUMAN	P52742 homo sapien

RESULT 1
 IKAR_HUMAN
 ID IKAR_HUMAN STANDARD; PRT; 519 AA.
 AC Q13422; Q00598;
 DT 15-DEC-1998 (REL. 37, CREATED)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1).
 GN IKAROS OR IK1 OR OR LYF1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=BONE MARROW;
 RX MEDLINE: 96252222.
 RA NIETSELD W., MEYERHANS A.;
 RT "Cloning and sequencing of hik-1, a cDNA encoding a human homologue
 of mouse Ikaros/Lyf-1.";
 RL IMMUNOL. LETT. 49:139-141(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96132984.
 RA MOLNAR A., WU P., LARGESPADA D.A., VORTKAMP A., SCHERER S.,
 COPELAND N.G., JENKINS N.A., BRUNS G., GEORGIOPOULOS K.;
 RT "The Ikaros gene encodes a family of lymphocyte-restricted zinc
 finger DNA binding proteins, highly conserved in human and mouse.";
 RL J. IMMUNOL. 156:585-592(1996).
 CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
 THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
 MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
 CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
 PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED
 DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMUS, SPLEEN AND
 PERIPHERAL BLOOD LEUKOCYTES AND LYMPH NODES. LOWER EXPRESSION IN
 BONE MARROW AND SMALL INTESTINE.
 CC -!- ALTERNATIVE PRODUCTS: SIX FORMS OF THE PROTEIN: ISOFORMS IK1-IK6;
 ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS
 THAT OF IK1.
 CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U40462; G1299371; -
 DR EMBL: S80876; G1911483; -
 DR MIM: 603023; -
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 DR PFAM: PF00096; zf-C2H2; 5.
 KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
 KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
 FT DOMAIN 117 224
 FT ZINC-FINGERS I.
 FT ZINC-FINGERS II.
 FT ZN_FING 117 139
 FT C2H2-TYPE.
 FT ZN_FING 145 167
 FT C2H2-TYPE.
 FT ZN_FING 173 195
 FT C2H2-TYPE.
 FT ZN_FING 201 224
 FT C2H2-TYPE.
 FT ZN_FING 462 484
 FT C2H2-TYPE.
 FT ZN_FING 490 514
 FT C2H2-TYPE.

THIS PAGE BLANK (USPTO)

Qy 448 -----DVFRCDHCHVLFLOYVMFTIHMGHGGERDPF 478
| | | | : | : | :
Db 393 SHTSPSTFEFGVTEDNHQSPSKDHTGKEKFFSCSECKSFFKVSLDKHLVVHTGEKPY 452

```

QY      479 ECNMG-----YRSHDR 490
      |  | |
Db      453 HCIEGRSYTHQSSLKSHQR 472

```

RESULT 15

S26823
finger protein ZNF43 - human
C:Species: Homo sapiens (man)
C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1998
C:Accession: S26823
R:Lovering, R.; Trowsdale, J.
Nucleic Acids Res. 19, 2921-2928, 1991
A:title: A gene encoding 22 highly related zinc fingers is expressed in lymphocytes
A:Reference number: S26823; MUID:91279444
A:Accession: S26823
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-803 <LOV>
A:Cross-references: EMBL:X59244; NID:g38031; PID:g38032
C:Genetics:
A:Gene: GDB:ZNF43; HTF6
A:Cross-references: GDB:128653
A:Map position: 19p13.1-19p12
C:Keywords: DNA binding

Query Match	10.8%	Score 295.5;	DB 2;	Length 803;
Best Local Similarity	37.3%;	Pred. NO. 1.1e-13;		
Matches 59; Conservative	29;	Mismatches 51;	Indels 19;	Gaps 3;

```

Qy 81 EESEMPYAREYSDYESIKLERHVPYD-----NSRPTSGKM-----NCDVC 122
    ||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 479 EECGKAFRSSNLTKHKKIHIIEK-PYKCECGKAFKWSSKLTETHKI'HTGKPYKCEEC 537

```

Oy 123 GLSCISFNVLMVHKRSHTGERPFQCNQCASFTQGNLLRIHLKHTGEKPFKCHLCNYAC 182
- : - :: | | | | | : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Dd 538 GKAFNHFSILTKHRIHTGSKPYKEECGKAFTQSSNLTTHKKIHTGKFYKCCECGKAF 597

Qy. 183 QRRDALTGHLRTHSVEKPYKCEFCGRSYQORSSLEEKK 220
 : ||| : | : ||||| ||:: | : | : ||
 Db 598 TQSSNLTHKKIHTGKGPKYKEECGKAFNOFSTLTTHK 635

Search completed: November 6, 1999, 08:08:51
Job time: 71 sec

QY 59 KVKDEYSDRDENIMKPEPM-----GDAESEMPPSYAREYSDYESIKLERHVPYD 108

Db 343 H-----LPCTECGGTFTNEQELLAHQSTHTEEQKPLP--CTECGEIFSDEHELLTHQ 392

Db 343 H-----LPCTECGGTFTNEQELL

Qy	142	ERPQCNOCGASFTQKNLRLHKLHTGKEKPFKCHLCNVAQORRDALTLGHLRTHSVKPY	201
Db	185	EKPYQCSHCSCFSQHSQGMVKHLRIHTGKPYMCNHCYKHFSSQSDLIKHQRIHTGKPY	244
Qy	202	KCEFCGRSYKORSLEEHKE-----RC-RAFLQNPDLGDAASVEARHIKAEMGS	249
Db	245	KCDVCGKRAFGSSDRILHQRIHTGKPYPCAQCNCKSFQNSDL-----IKHRRILHTEG	297
Qy	250	ERALVLDRLASNAVKRSSMPQKFIGEKHRCFDANTNPGYMYKEKEMQMTRMDQAIN	309
Db	298	K-----PYKSECCKAFNQSSVLILHQRIHTGKPY-PCNOCTKS-----FSRLSD-LINH	346
Qy	310	AISYLGAERFPLVQTPATSEMPVSISSVYPIALTRADPM-----GAPOHEK-----	361
Db	347	QRIHTGK-----PYCSCQCKMFS-----RRSDLVKHYRIHTGKPYECDKCG	390
Qy	361	-----KRILLPEKLPSEKGLSPNNSAQD-----S-----S	385
Db	391	KTFSSQSNLILHQRIHTGKPYPCNCSCKSFSGSDLIKHQHVHTGKPYTCNLCSKSF	450
Qy	386	TDITSDNHEDRQH-----LYQOSHVL-----POARNGMPLLEKVPKS	422
Db	451	QSSDLTKHQHVHSGEKPYHCSSCNKAFRQSSDLILHRVHTGSRPYA-----CTQCPRS	504
Qy	423	F-----ELLKPPPPICLRDSIKVINKEGVMDFVRCDHCVLFLDYVMTFHMGCGRDFP	478
Db	505	FSQKSDLIKHQRI-----HTGE-KPYKC-MCGKRAFSQCSAFTLHQRIHTGKPY	551
Qy	479	ECNWCGRYSRHDREYFSSH	496
Db	552	PCAQCCKGSFQSRSDLVNH	569

RESULT 10
B34612
zinc finger protein ZNF8 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Jun-1990 #sequence_revision 09-Oct-1992 #text_change 14-Aug-1998
C:Accession: B34612
R:Lania, L.; Doniti, E.; Pannuti, A.; Pascucci, A.; Pengue, G.; Feliciello, Genomics 6, 333-340, 1990
A:Title: cDNA isolation, expression analysis, and chromosomal localization
A:Reference number: A34612; MUID:90169993
A:Accession: B34612
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-543 <LAN>
A:Cross-references: GB:M29581; NID:G340447; PID:G340448; GB:J04751
C:Genetics:
A:Gene: ZNF8
A:Cross-references: GDB:120510; OMIM:194532
A:Map position: 20q13-20q13
C:Keywords: DNA binding; zinc finger

```

Query Match      10.9%  Score 298.5;  DB 2;  Length 543;
Best Local Similarity 29.9%;  Pred. No. 4.1e-14;
Matches 79;  Conservative 40;  Mismatches 80;  Indels 65;  Gaps 9;

QY 12 STEEQPLTFESP-----DALNDYSLPFPKPEHEN-----VDSREA 45
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 75 SRKEEGLPEEPESHVTRGEGPTDAPYPTTLGKDRECCQSALKEQNKLQLEFGLEA 134
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 46 PANEDEDAGDSMKVKDEYSRDEINIM--KPEPMGDAEESEMPYSYAREYSDYE--SI 99
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 PV---QDOAYKTLRL-----ENCVLSSGFNPFPESRGEYLTYDSQITDSEHNSL 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 100 KLER-----HVPYDONS-----RTSGKM-----NCDVCGLSCTSFENVLMVHK 136
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 VSQQTGSPGKPGENSCHDRSSQAIPTEUTKSOVDKPKYKCTDCGKSFHNNAHLTVHK 244
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 137 RSHTGERPFCQNOCGASPTQGNLLRIKHLKLTGEKPFKCHLCNTACQRRDALTGHLRTHS 196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      245  RIHTGERPYMKCEGKAFSONSSLVQHERIHTGDKPKYKCAEGKSFCHSTHLTVHRRINT 304
Qy      197  VEKPKYCEFGCRSYKORSSLEEKK 220
          |||||:||||:|||||
Db      305  GEKPYEQDCGRAFNQSSSLGRHK 328

RESULT  11
finger protein (clone xlcOP28) - African clawed frog (fragment)
C:Species: xenopus laevis (African clawed frog)
C:Date: 28-Feb-1990 *sequence_revision 28-Feb-1990 *text_change 31-Dec-1999
C:Accession: S06556
J. Netfield, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poetli
R. Mol. Biol. 208, 639-659, 1989
A:Title: Second-order repeats in Xenopus laevis finger proteins.
A:Reference number: S05632; MUID:90040698
A:Accession: S06556
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-439 <NIE>
C:Keywords: DNA binding; zinc finger

Query Match          10.9%; Score 298; DB 2; Length 439;
Best Local Similarity 24.7%; Pred. NO. 3.4e-14;
Matches 98; Conservative 53; Mismatches 136; Indels 110; Gaps

Qy      96  YESIKLERHPYDINSRPTSGKMNCMDYGLSCISFNVLVYHKRSHTGERPFCNQCASFT 155
          |||||:||||:|||||
Db      129  YPSNLVQVHRTHTGDRP-----FOCTECDKSFYTKMSKLVMLRIHTGKPKYKCECDKSPS 184

```

[illegible]

RESULT . 12
G02075
transcription repressor zinc finger protein 85 - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
C:Accession: G02075
R:Poncelet, D.A.
submitted to the EMBL Data Library, September 1995
A:Reference number: G09169
A:Accession: G02075
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-595 <PON>
A:Cross-references: EMBL:U35376; NID:g1017721; PID:g1017722
C:Genetics:
A:Gene: GDB:ZNF85

QY 224 RAFLQNPDLGDAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFIKGRHCFDA 283
Db 170 -----DSAN----- 174
QY 284 NYNPGYMEKENEMOTRMMDQAINNAISYLGAERPLVQTPAPTSEMVPVISVYPI 343
Db 174 -----YEKE-DMTSHVMDQAINNAIYLGAESLRPLVQTPPG-SSEVPVYSMWQL 224
QY 344 ALTRADMP-----MGAPQEMKRIILPEKILPSERGLSPNNSAQDSTDSDNHEDRQH-- 398
Db 225 HKPPSDGPPRNSHSAQDAVDNLLLSKAKSVSREASPSNCSQDSTDDESNAEQRSL 284
QY 398 LYQOSHVLVPOARNGMPLKVPKRSFELLKPPICLRDSIKVINKEGEVMDVFRDCHV 457
Db 285 IYLTNH-INPHARNGI-ALKEQRAYEVLRAASNSQDAFRVYSTGEOLKYKKEHCRV 342
QY 458 FLFDYVMFTIHMCHGFRDPFCNMGCGYRSHDRYEFSSHIARGEHR 503
Db 343 LFLDHWYTHMGCHGFRDPFCNMGCGYRSHDRYEFSSHIARGEHR 388
RESULT 5
S22954
finger protein zfp-37 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
C:Accession: S22954
R:Burke, P.S.; Wolgemuth, D.J.
Nucleic Acids Res. 20, 2827-2834, 1992
A:Title: Zfp-37, a new murine zinc finger encoding gene, is expressed in a developmental
A:Reference number: S22954; MUID:93210982
A:Accession: S22954
A:Molecule type: mRNA
A:Residues: 1-553 <BUR>
A:Cross-references: EMBL:X64413; NID:g55474; PID:g55475
C:Superfamily: zinc finger protein ZFP-36
C:Keywords: DNA binding

Query Match 12.6%; Score 344; DB 2; Length 553;
Best Local Similarity 23.3%; Pred. No. 2,3e-17;
Matches 114; Conservative 55; Mismatches 129; Indels 192; Gaps 16;

QY 21 ESPDALNDYSLPKPHIEINVDREAPA-----NEDEAGEDSMKVDEYSRDENI 71
Db 121 KPPTANEHRKSLSHSASDVNDIEPTRRKKCDKLPNNKLSRKG-----KNOTSKKCKV 175
QY 72 MKPEPMGDAE-----SEMPYSYAREYSDYESIKLERHVPYDNRPTSGK-MNCDVCGLS 126
Db 176 CHRSHATKEDKIOTGEKKRSHCRTPSKPE-----KAPGSGKPYECNHCQKVL 223
QY 127 ISFNVLVHKSHTGERPQCQCGASFTQKGNLLRHILKHTGEKPFKCHLCNACQRRD 186
Db 224 SHKQGLLDHQRHTGTEKPYECNECGIAFSQKSHLVHQRHTGTEKPYECQCGKAHGHK 283
QY 187 ALTGHLRTHSVKPYKCFGRSYKORSLEEH-----KERCRAFLQNPDLGD 234
Db 284 ALTDHLRIHTGTEKPYKCNCEGTFRSHSNLMQLRSHTGTEKPYECKGCKSPRYNSSL-- 342
QY 235 AASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFIKGRHCFDANYPGYMEKE 294
Db 342 -----TEHVRTHG-----EIPY-----ECNEC----- 360
QY 295 NEMOTRMMDQAINNAISYLGAERPLVQTPAPTSEMVPVISVYPIALTRADMPGA 354
Db 360 -----GKAFK-----YGSLLTK----- 372
QY 355 POEMKKRILLPEKILPSERGLSPNNSAQDSTDSDNHEDRQHLVQOSHVLVPOARNGM 414
Db 372 -----HMRHTGK-----PFE-----CNECGKTFKSHLVHQRTH-- 405
QY 415 LLKVPKRSFELLKPPICLRDSIKVINKEGEVMDVFRDCHVFLDYVMFTIHMCHGCF 474

Db 405 -TKEXP-----YKDECGKAFGHSSSLTYHMRHTG 434
QY 475 RDPFECNMG 484
Db 435 DCPFECNCG 444
RESULT 6
S10245
finger protein, testis - mouse
C:Species: Mus musculus (house mouse)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 10-Sep-1997
C:Accession: S10245
R:Neikii, D.; Dudley, K.; Cunningham, P.; Akhavan, M.
Nucleic Acids Res. 18, 3655, 1990
A:Title: Cloning and sequencing of a zinc finger cDNA expressed in mouse testis.
A:Reference number: S10245; MUID:90301500
A:Accession: S10245
A:Molecule type: mRNA
A:Residues: 1-411 <NEL>
A:Cross-references: EMBL:X52533; NID:g53456; PID:g53457
C:Keywords: DNA binding; zinc finger

Query Match 11.7%; Score 319; DB 2; Length 411;
Best Local Similarity 24.6%; Pred. No. 9,8e-16;
Matches 95; Conservative 41; Mismatches 88; Indels 162; Gaps 12;

QY 112 PTSGK-MNCDVCGLSISFNVLVHKSHTGERPFOCQCGASFTQKGNLLRHILKHTGE 170
Db 66 PGGKPYECNHCQGVLSHKQGLLDHQRHTGTEKPYECNECGIAFSQKSHLVHQRHTG 125
QY 171 KPFKCHLCNACQRRDALTGHLRTHSVKPYKCFGRSYKORSLEEH----- 220
Db 126 KPYECQCGKAHGHKHALTDHLRIHTGTEKPYKCNCEGTFRSHSNLMQLRSHTGTEKPYE 185
QY 220 -KERCRAFLQNPDLGDAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFIKGR 278
Db 186 CKEGCKSPRYNSSL-----TEHVRTHG-----EIPY-----ECN 215
QY 279 HCFDANYPGYMEKENEMOTRMMDQAINNAISYLGAERPLVQTPAPTSEMVPVIS 338
Db 216 EC-----GKAFK----- 223
QY 339 SVYPIALTRADMPGAPOEMKKRILLPEKILPSERGLSPNNSAQDSTDSDNHEDRQHL 398
Db 223 -YGSLLTK-----HMRHTGK-----PFE-----CNECGKTF 249
QY 399 YQOSHVLVPOARNGMPLKVPKRSFELLKPPICLRDSIKVINKEGEVMDVFRDCHV 458
Db 250 SKKSHLVHQRTH--TKEXP-----YKDECGKA 276
QY 459 FLDYVMFTIHMCHGFRDPFECNMG 484
Db 277 FGHSSSLTYHMRHTGDCPFECNCG 302
RESULT 7
S00647
finger protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 31-Dec-1993
C:Accession: S00647
R:Ruiz i Altaba, A.; Perry-O'Keefe, H.; Melton, D.A.
EMBO J. 6, 3065-3070, 1987
A:Title: Xfin: an embryonic gene encoding a multifingered protein in Xenopus.
A:Reference number: S00647; MUID:88082679
A:Accession: S00647
A:Molecule type: mRNA
A:Residues: 1-1350 <RUI>
A:Cross-references: EMBL:X06021
A:Note: It is uncertain whether Met-1, Met-11, Met-16, Met-38, or Met-39 is the Initi
C:Genetics:

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C:Accession: I59572
R:Georgopoulos, K.; Moore, D.D.; Delfler, B.
Science 258, 808-812, 1992
A:Title: Ikaros, an early lymphoid-specific transcription factor and a putative mediator of hematopoietic differentiation
A:Reference number: I59572; MUID:93068267
A:Accession: I59572
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-431 <RES>
A:Cross-references: GB:L03547; NID:g198286; PID:g198287
C:Genetics:
A:Gene: Ikaros

Query Match 42.4%; Score 1156; DB 2; Length 431;
Best Local Similarity 51.3%; Pred. No. 9.2e-76;
Matches 245; Conservative 64; Mismatches 93; Indels 76; Gaps 15;

```
QY 48 NEDED-----AGDSMKVKDEYSDRDENIMKPEPMGDASESEMPYSIYAREYSYIKLER 103
D 4 DEGDMSQVSGKESPPVSDTPDEGDEPMPVPEDLSTTSGAQ----- 45

QY 104 HVPYDNSRPTSGKMNCDCGLSCISFNVLVHVKRSHTGGERPQCNOCGASFTQKGNLLRH 163
D 45 -----QNSKSDRG-----MGERPQCNOCGASFTQKGNLLRH 76

QY 164 IKLHTEKPFKCHLCNYACORRDALTGLRTHSVKPKYKCEFCGRSYKORSLEEHKERC 223
D 77 IKLHSEKPFKCHLCNYACRRRDALTGLRTHSVGPKHCKGCGYKORSLEEHKERC 136

QY 224 RAFLNQ---POLGDAASVEARHIK-----AEMGSRALVLDRLASNVAKRKSSMPQKFTG 275
D 137 HNYLESMLGPGVPCVKEETNHENAEADLCKIGAERSLVLDRLASNVAKRKSSMPQKFTG 196

QY 276 EKRRHF--LYQSHVVLPOARNGMPLLEKVPKPRFELLKPPICLRDSTIKVINKGEVMD 448
D 311 NAEQSRGLIYLTNH--INPHARNGL--ALKEQRAYEVLRAASENSQDAFRVYSTSGEQLK 369

QY 449 VPRCDHCHVFLDYVMFTTHM--GCHGFRDPFCNMGCGYRSHDRYEFSSHTARGHR 503
D 370 VYKCEHCRVFLDHVMYTHMGCHGCHGFRDPFCNMGCGYRSHDRYEFSSHTARGHR 427
```

RESULT 3
B56229
Lymphoid transcription factor Ikaros/Lyf-1, form V - mouse
C:Species: Mus musculus (house mouse)
C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
C:Accession: B56229
R:Hahn, K.; Ernst, P.; Lo, K.; Kim, G.S.; Turck, C.; Smale, S.T.
Mol. Cell. Biol. 14, 7111-7123, 1994
A:Title: The lymphoid transcription factor Lyf-1 is encoded by specific, alternatively spliced transcripts
A:Reference number: A56229; MUID:95021239
A:Accession: B56229
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-427 <HAH>
A:Cross-references: GB:S74518
C:Keywords: alternative splicing

Query Match 42.3%; Score 1154; DB 2; Length 427;
Best Local Similarity 50.9%; Pred. No. 1.3e-75;
Matches 242; Conservative 65; Mismatches 94; Indels 74; Gaps 14;

QY 48 NEDED-----AGDSMKVKDEYSDRDENIMKPEPMGDASESEMPYSIYAREYSYIKLER 103
D 4 DEGDMSQVSGKESPPVSDTPDEGDEPMPVPEDLSTTSGAQ----- 45
QY 104 HVPYDNSRPTSGKMNCDCGLSCISFNVLVHVKRSHTGGERPQCNOCGASFTQKGNLLRH 163
D 45 -----QNSKSDRG-----MGERPQCNOCGASFTQKGNLLRH 76
QY 164 IKLHTEKPFKCHLCNYACORRDALTGLRTHSVKPKYKCEFCGRSYKORSLEEHKERC 223
D 77 IKLHSEKPFKCHLCNYACRRRDALTGLRTHSVGPKHCKGCGYKORSLEEHKERC 136
QY 224 RAFLNQ---POLGDAASVEARHIK-----AEMGSRALVLDRLASNVAKRKSSMPQKFTG 275
D 137 HNYLESMLGPGVPCVKEETNHENAEADLCKIGAERSLVLDRLASNVAKRKSSMPQKFTG 196
QY 276 EKRRHF--LYQSHVVLPOARNGMPLLEKVPKPRFELLKPPICLRDSTIKVINKGEVMD 448
D 311 NAEQSRGLIYLTNH--INPHARNGL--ALKEQRAYEVLRAASENSQDAFRVYSTSGEQLK 368
QY 449 VPRCDHCHVFLDYVMFTTHM--GCHGFRDPFCNMGCGYRSHDRYEFSSHTARGHR 503
D 369 VYKCEHCRVFLDHVMYTHMGCHGCHGFRDPFCNMGCGYRSHDRYEFSSHTARGHR 423
RESULT 4
A56229
Lymphoid transcription factor Ikaros/Lyf-1, form IV - mouse
C:Species: Mus musculus (house mouse)
C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
C:Accession: A56229
R:Hahn, K.; Ernst, P.; Lo, K.; Kim, G.S.; Turck, C.; Smale, S.T.
Mol. Cell. Biol. 14, 7111-7123, 1994
A:Title: The lymphoid transcription factor Lyf-1 is encoded by specific, alternatively spliced transcripts
A:Reference number: A56229; MUID:95021239
A:Accession: A56229
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-392 <HAH>
A:Cross-references: GB:S74517
C:Keywords: alternative splicing

Query Match 24.3%; Score 664; DB 2; Length 392;
Best Local Similarity 32.7%; Pred. No. 1.6e-40;
Matches 172; Conservative 58; Mismatches 112; Indels 184; Gaps 15;

```
QY 12 STEEPLTEPDALNDYSLPKPHEIENVDSREAPANED-----EDA 53
D 13 SKESPPVSDTDE--GDEPMPVPEDLSTTSGAQNSKSDRGVAYGADGFRDFHAIISDR 71
QY 54 GDSMKVKDEYSDRDENIMKPEPMG-----DAESEMPYSIYAREYSYIKLER 103
D 72 GMPG--NVKVTQSDENGRCACEMNGEAEADLRLMDLDSGCKMNGSHRDQGSALS----- 126
QY 104 HVPYDNSRPTSGKMNCDCGLSCISFNVLVHVKRSHTGGERPQCNOCGASFTQKGNLLRH 163
D 126 --GVGIGRLPNGLKCDICGIVICGNVLMVHVKRSHTD----- 162
QY 164 IKLHTEKPFKCHLCNYACORRDALTGLRTHSVKPKYKCEFCGRSYKORSLEEHKERC 223
D 162 -----KC-----LSDMPY----- 170
```

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 08:07:40 ; Search time 18.18 Seconds
(without alignments)
1117.338 Million cell updates/sec

Title: US-09-019-348-2
Perfect score: 2728
Sequence: 1 MEDIQTVELKSTEEQLPT.....HDIYFSSHTARGEHRAMLK 507

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR_60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1286	47.1	518	2	A56355	DNA-binding protei
2	1156	42.4	431	2	I59572	Ikaros DNA binding
3	1154	42.3	427	2	B56229	lymphoid transcrip
4	664	24.3	392	2	A56229	lymphoid transcrip
5	344	12.6	553	2	S22954	finger protein zip
6	319	11.7	411	2	S10245	finger protein, te
7	316	11.6	1350	2	S00647	finger protein - A
8	309	11.3	636	2	I48689	gene NK10 protein
9	307	11.3	580	2	A37107	spermatogenesis pr
10	298.5	10.9	543	2	B34612	zinc finger protei
11	298	10.9	439	2	S06536	finger protein (cl
12	297.5	10.9	595	2	G02075	transcription repr
13	296.5	10.9	510	2	A36901	chorion transcript
14	295.5	10.8	675	2	S51037	zinc-finger protei
15	295.5	10.8	803	2	S26823	finger protein ZNF
16	295	10.8	693	2	I37570	zinc finger protei
17	293.5	10.8	485	2	A40751	finger protein ZNF
18	293.5	10.8	393	2	JN0533	finger protein pML
19	293.5	10.8	614	2	JH0500	zinc finger protei
20	293	10.7	295	2	A46017	zinc finger protei
21	292.5	10.7	1191	2	S35305	zinc finger protei
22	292	10.7	378	2	S00549	developmental cont
23	291	10.7	378	2	S33994	finger protein ZNF
24	289.5	10.6	491	2	S00753	finger protein HF
25	288.5	10.6	519	2	A38073	transcription acti
26	288	10.6	545	2	I49636	DNA-binding protei
27	286.5	10.5	651	2	B32891	finger protein 2,
28	286.5	10.5	292	2	S43826	finger protein OZF
29	286	10.5	428	2	S03677	finger protein (cl
30	284.5	10.4	247	2	S47070	finger protein HZF
31	283.5	10.4	591	2	S65088	finger protein XFO
32	283.5	10.4	427	2	A35659	krueppel-related p
33	282	10.3	474	2	I54338	zinc finger protei
34	281	10.3	247	2	S06553	finger protein (cl
35	279	10.2	732	2	S47073	finger protein HZF
36	277.5	10.2	671	2	JE0288	krueppel-type zinc
37	277.5	10.2	542	2	A54661	zinc finger protei
38	277	10.2	428	2	A32891	finger protein 1,
39	277	10.2	469	2	I38600	zinc finger protei

40 276.5 10.1 462 2 S10397 finger protein kox
41 276 10.1 453 2 S06544 finger protein (cl
42 276 10.1 353 2 S65086 finger protein XFO
43 276 10.1 348 2 I38599 zinc finger protei
44 276 10.1 794 2 S59069 z13 protein - mous
45 275 10.1 576 2 A48157 renal transcriptio

ALIGNMENTS

RESULT 1
A56355
DNA-binding protein Ikaros form 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C;Accession: A56355
R;Molnar, A.; Georgopoulos, K.
Mol Cell Biol. 14, 8292-8303, 1994
A;Title: The Ikaros gene encodes a family of functionally diverse zinc finger DNA-bin
A;Reference number: A56355; MUID:95059058
A;Accession: A56355
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-518 <MOL>
C;Keywords: alternative splicing; DNA binding

Query Match 47.1%; Score 1286; DB 2; Length 518;

Best Local Similarity 53.1%; Pred. No. 5.5e-85;

Matches 276; Conservative 69; Mismatches 129; Indels 46; Gaps 17;

Qy 12 STEEQPLTSPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVKDEYSDRDENI 71
Db 13 SGKESPPVSDTPDE-GDEPMVPVEDLSTTSG--AQQNSKSDRGWAS-NVKVETQSDENG 68
Qy 72 MKPEPMG-----DAESEMPYAYAREYSDYESIKLERHVPYDINSRSTSGKMNCDV 121
Db 69 RACEMNGEECAEDLUMILDASGEKMGNSHRDQGSALS-----GVGGIRLPNGKLKCDI 121
Qy 122 CGLSCIGFNVLVHKRSHTERPPQCNOCASFTQKGNLLRHKLHTGERPKFCHLCNYA 181
Db 122 CGIVCIGPNVLVHKRSHTERPPQCNOCASFTQKGNLLRHKLHTGERPKFCHLCNYA 181
Qy 182 QQRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEKERKCRFLQN---PDLDGAASV 238
Db 182 CRRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEKERKCHNYLESMLPGMYPVIKE 241
Qy 239 EARHIK-----AEMGSEALVLDRLASNVAKRKSSMPQKFIGEKHCF-DANYNPGYME 292
Db 242 ETNENMAEDLCKTGASERSLVDRLASNVAKRKSSMPQKFIGDK--CLSDMPYDSA-NYE 298
Qy 293 KENEMOTRMMDQAINNAISVLGAERPLVQTPPTSEMPVVISSVPIALTRADMP- 352
Db 299 KE-DMTSHVMDQAINNAISVLGAERPLVQTPPG-SSEVPVVISSMYQLHKPPSGPP 356
Qy 352 ---MGAPQEMEKRIKLPSELKIPSERGLSPNNSAQDSTDSDSNHEDRQH---LYQQSHVYL 406
Db 357 RSNHSAQDAVDNLLLSKAKSVSSEREASPSNSQDSTDSDTESNAEQRSGLIYLTNH-IN 415
Qy 407 POARGHPLKVEPRSFELKPPPICLRDSIKVKNGEVDVFCRDCHVFLDYVMT 466
Db 416 PHARNGL-ALKEQRAYEVLRAASENSQDAFRVYSTGSEQLKYKCEHCRVFLDHYMT 474
Qy 467 IHM---GCHGFRDPFECNMGYRSHDRYFSSHTARGEHR 503
Db 475 IHMCHGCHGFRDPFECNMGYHSQDRYFSSHTIRGEHR 514

RESULT 2
I59572
Ikaros DNA binding protein - mouse
C;Species: Mus musculus (house mouse)

THIS PAGE BLANK (USPTO)

[illegible]

Search completed: November 6, 1999, 09:41:35
Job time: 67 sec

Db 193 SLESLVVEPQVPEASGAPDIYSSNKDLPPETPGAEDKOKSSKTKPFROCKPC-QYEAASE 251
QY 49 ED-----EDAGDSMKVKDEYSDRDENIMKPEPMGDAESEM---PYSYA 90
Db 252 EQFVHHIRVHSAKKFFVEESAQAARSGSSTAEE-----GDFSGKPIRCDRCGYN 304
QY 91 REYSDYESIKLERHV-PYONSRTPSKMCNDCVGLSCIS-----129
Db 305 TNRDHYTALHKKHHTRAGNER---VYKCIITTYTVSEYHWRKHLRHHFPRKVVTCGK 360
QY 129 -----FNVLMVHKSHSGTGERPFOCNOCGAGSTOKGNLLRHKLHTGKPKFCHLCNTVA 181
Db 361 CNYFSDRKNNYVQVHTHTGERPKYKELCPYSSQTHLTRHWRTSGEKKPKFCDCQSTV 420
QY 182 CORRDLATGHLR-THSVEKPYKCEFCGRSYKQSSLEHKE---RCRAFLOPNPDGLDAAS 237
Db 421 ASNOHEVTRHAROVHNGPKPLNCPHCDYKTAQRSNEKKHVELHVNPRQF-NCPVCDYAAS 479
QY 238 VEAR---HIKAEMGS--ERALVLDRL-ASNVAKRKSSMPQKFTGEGKRHCDFDANYNPGYMY 291
Db 480 KKNLQYHFKSKHPTCPNTMDVSKYKLLKTKKREADLPDNTINEK-----526
QY 292 EKENEMQMTRMM-DOAINNAISYLGAEAFRPLVQTPAPTSEMVPVVISSVPIALTRADM 350
Db 526 ---TEIEQTKIKGDVAGKKNKSVKAEKRDVSKKKPSNNVSVIQVTTTRKSVTEVKEM 582
QY 351 PMGAPOEMKKRILLPEKILPSRGLSPNNSAODSTDTSNH---EDROHLYQOSSHVILP 407
Db 583 DVHTGNSSEKF-----SKTKSKRKL-----EVDSHSLGHPVNDSESTKKKKKVES 629
QY 408 OARNGMPLLEKVRP---SPELLKPPICLRDSTK 438
Db 630 KSKNS---QEVPKGDSKVEENKQNTCMKKSTK 660

RESULT 14

US-08-718-661-2
; Sequence 2, Application US/08718661
; Patent No. 5876972
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
; TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
; NUMBER OF SEQUENCES: 15
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,661
; FILING DATE:
; CLASSIFICATION: 530
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 667 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-718-661-2

Query Match 8.2%; Score 224.5; DB 2; Length 667;
Best Local Similarity 21.7%; Pred. No. 6e-14;
Matches 89; Conservative 52; Mismatches 166; Indels 103; Gaps 14;

QY 119 CDVCGLSGISFNVLWYHKRSHHTGERPFOCN--QCGASFTQKGNLLRHKLHTGKPKFCH 176
Db 6 CQCKGSGFTVLEKFTLHNSHRSRERFKCSKAECGAFVSKYKLMRHMATHSPQKHQCT 65
QY 177 LCNYACQRDLATGHLRTHSVKE-PYKCFEGCGRSYKQSSLEEHKRCRAFLQNPDLGDA 235
Db 66 HCEKTNRKDLKHLNHLQTHDPNKAISTACDDCGKKYHMLGYKHH-----LALHSASNGD- 120

QY 236 ASVEARHKAEMGSRALVLDRLASNVAKRKSSMP-----271
Db 120 --LTCGVCTLEIGSEVL-LDHLKSHAEKANQAAPREKKYQCDCHDCRCFYTRKDYRRHLV 176
QY 271 -----OKFTGKRRHC-----FDANYN---PGYMYE 292
Db 177 VHTGCKDFLCQCAQRF-GRKDHLTRHTKTHSQELMOENMQAGDYQSNFQLIAFSTSFQ 235
QY 293 KENEMQMTRMMDOAINNAIS-----YLGAEAFRPLVQTPAPTSEMVPV---I 337
Db 236 IKVDMPMPFOLGAAPENGLDGLPPEVHGLVLAARPEAPQPMPPLEPLEPLEPLEPM 295
QY 338 SSVYPIALTRADMPGAPQEMEKRIILP-----EKILPSE---RGLSPNNSAODSTDTSN 391
Db 296 QSLEPLQLPLEMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPM 355
QY 392 HEDROHLYQOSSHVILPQARNGMPLLEKVRPSEL-----LKPPPICLRD 435
Db 356 QPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPM 405

RESULT 15

US-08-040-548-1
; Sequence 1, Application US/08040548
; Patent No. 5763209
; GENERAL INFORMATION:
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
; TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/040,548
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coughlin, Daniel F.
; REGISTRATION NUMBER: 36,111
; REFERENCE/DOCKET NUMBER: arcd067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 744-0090
; TELEFAX: (312) 245-4961
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-040-548-1

Query Match 8.0%; Score 218.5; DB 2; Length 533;
Best Local Similarity 26.4%; Pred. No. 1.7e-13;
Matches 71; Conservative 23; Mismatches 80; Indels 95; Gaps 10;

QY 142 ERPFQ--NQCGASFTQKGNLLRHKLHTGKPKFCHLCNYACQRDLATGHLRTHSVKE 199
Db 333 ERPYACPVESCDRRFSRSDLTRHRIHTGQKPFQCRICMRNFSRSDHLTTHRTHTGCK 392
QY 200 PYKCFEGCGRSYKQSSLEEHKRCRAFLQNPDLGDAASVEARHKAEMGSRALVLDRLA 259

MOLECULE TYPE: protein
PCT-US94-06669-2

Query Match 9.7%; Score 265.5; DB 3; Length 706;

Best Local Similarity 27.8%; Pred. No. 5.3e-18;

Matches 70; Conservative 39; Mismatches 88; Indels 55; Gaps 9;

5 QPVVEIKSTEQPLPTSPDALNDYSLPKPHEINVDRE---APANEDE-----52

Db 415 QPPMEPNLDIQS-PTKLSASGSDTIPQASRLNIVNRSMGTGSPRSSSESHPLYMHPP 473

52 --DAGDSMKVRDEYSDROENINKP---BPMGDAESEMPSYAREYSD-----96

Db 474 KTSCGSGQS---POHAEMCLHTAGPTFAEMGETQS-----EYSSSSENGAFFC 520

QY 96 -----YESIKLRRHPVYDNS-RPTSGKMNCDVGLSCISFNVLMVHKRSHTGERPFOC 147

521 NECDCRFSEASLKRHTLQTHSDKP---YKDCRQASFRYKGNLASHKTVHTGKPYRC 576

QY 148 NOCGASTQKGNLRLHKLHTGEPFKCHLCNYACQRRDALTGHLRTHSVKPKYKCFGC 207

Db 577 NICGAQNRPANLXTHRIHSGERPKYCETCGARFVQVAHLRAHVLHTGKPYPCBICG 636

QY 208 RSYKORSLSLEH 219

Db 637 TRFRHLQTLKSH 648

RESULT 12

PCT-US95-08429-5

Sequence 5, Application PC/TUS9508429

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: CTCF

NUMBER OF SEQUENCES: 21

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/08429

FILING DATE: 15-JUN-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/261,680

FILING DATE: 17-JUN-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14538A-11-IPC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-467-9600

TELEFAX: 415-543-5043

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 728 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-08429-5

Query Match 9.6%; Score 262.5; DB 3; Length 728;

Best Local Similarity 25.7%; Pred. No. 1.1e-17;

Matches 78; Conservative 44; Mismatches 111; Indels 71; Gaps 11;

QY 9 ELKSTEEQPL-PTESPALN--DYSLPKPHIEINVDREAPANEDEAGEDSMKVKDEYS 65

Db 174 EVETLEQGLQPDQPNQWQPDYQPPAKTKKNNKSLRYTEGKDV---DVSVDYFEE 230

QY 66 DRDENI-----MKPEPMGDAESEMPSYAREYSDY---ESIKLERHV-PYDN 109

Db 231 EOEGLLSEVNAEKVVGNMKPPKPTIKKKGVKKTFCCLCYTCPRRSNLDHRMKSHTD 290

QY 110 SRPTSGKMNCDVGLSCISFNVLMVHKRSHTGERPFOCNOCGASTQKGNLL-----162

Db 291 ERP-----HKCHLCGRAFRIVTLRLNHLNHTGTGRPHKCPDCDMAFVTSGLVHRRYKHT 346

QY 162 -----RHKLHTGKPKFKCHLCNYACQRRDALTGHLRTHSVKPK 200

Db 347 HEKPFKCSMCDYASVEVSKLKRHIRSHNTGERPFOCSLCSYASRDTYKLRHRMTHSGEKP 406

QY 201 YKCFEGGRYKQRSS-----LEEKRCRAF-----LQNPDLGDAASVEARIKAE 246

Db 407 YECYICHAFTOSGTGMKHILOKHNTENVAKFCHPHCDTVIARSKDLG--VHLRKQHSYIE 464

QY 247 MGSE 250

Db 465 QGKK 468

RESULT 13

US-08-398-590A-40

Sequence 40, Application US/08398590A

Patent No. 5935811

GENERAL INFORMATION:

APPLICANT: Anderson, David J.

APPLICANT: Schoenherr, Christopher J.

TITLE OF INVENTION: Neuron-Restrictive Silencer Factor

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hombach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/398,590A

FILING DATE: 03-MAR-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/103,445

FILING DATE: 06-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: Silva, Robin M.

REGISTRATION NUMBER: 38,304

REFERENCE/DOCKET NUMBER: A-60897/RFT/RMS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

INFORMATION FOR SEQ ID NO: 40:

SEQUENCE CHARACTERISTICS:

LENGTH: 676 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-398-590A-40

Query Match 8.7%; Score 237; DB 2; Length 676;

Best Local Similarity 22.4%; Pred. No. 3.5e-15;

Matches 115; Conservative 68; Mismatches 203; Indels 128; Gaps 22;

QY 7 TVELKSTEEQP-----LPTESPDA--LNDYSLPKPHEINVDREAPAN 48

Db 7 TVELKSTEEQP-----LPTESPDA--LNDYSLPKPHEINVDREAPAN 48

; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-074-967-2

Query Match 9.7%; Score 265.5; DB 1; Length 706;
Best Local Similarity 27.8%; Pred. No. 5.3e-18;
Matches 70; Conservative 39; Mismatches 88; Indels 55; Gaps 9;

QY 5 OPTVELKSTEEQPLTPESDNDYSLPKPHEIENVDSRE---APANEDE----- 52
Db 415 QPPMEPNLDLQS-PTKLSASGEDSTIPQASRLNNIVNRMTGSPRSSSESHSPLYMHPP 473

QY 52 ---DAGEDSMKVDEYSDRDENIMKP---EPMGDAESEMPPSYAREYSD----- 96
Db 474 KCTSCGSQS---PQAEMLHTAGTFAEMGETQS-----EYSDSCENGAFCC 520

QY 96 -----YESIKLERHVPYDNS-RPTSGKMNCDVCGLSICISFNVLVHVKRSHGTGERPFC 147
Db 521 NECDCRFSEASLKRHTLQTHSDKP---YKCDRCQASFRYKGNLASHKTVHTGKPYRC 576

QY 148 NOCGASFTQKGNLRLHKLHTGKPKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCG 207
Db 577 NICGAQFNRPANLKTHTRIHSGKPKYKCEFCG---YKCDRCQASFRYKGNLASHKTVHTGKPYRC 576

QY 208 RSYKQSSLEEH 219
Db 637 TRFRHLQTLKSH 648

RESULT 10
US-08-553-541B-2
; Sequence 2, Application US/08553541B
; Patent No. 5882858
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Chaganti, Raju S.K.
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
; TITLE OF INVENTION: bcl-6
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,541B
; FILING DATE: May 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-553-541B-2

Query Match 9.7%; Score 265.5; DB 2; Length 706;
Best Local Similarity 27.8%; Pred. No. 5.3e-18;
Matches 70; Conservative 39; Mismatches 88; Indels 55; Gaps 9;

QY 5 OPTVELKSTEEQPLTPESDNDYSLPKPHEIENVDSRE---APANEDE----- 52
Db 415 QPPMEPNLDLQS-PTKLSASGEDSTIPQASRLNNIVNRMTGSPRSSSESHSPLYMHPP 473

QY 52 ---DAGEDSMKVDEYSDRDENIMKP---EPMGDAESEMPPSYAREYSD----- 96
Db 474 KCTSCGSQS---PQAEMLHTAGTFAEMGETQS-----EYSDSCENGAFCC 520

QY 96 -----YESIKLERHVPYDNS-RPTSGKMNCDVCGLSICISFNVLVHVKRSHGTGERPFC 147
Db 521 NECDCRFSEASLKRHTLQTHSDKP---YKCDRCQASFRYKGNLASHKTVHTGKPYRC 576

QY 148 NOCGASFTQKGNLRLHKLHTGKPKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCG 207
Db 577 NICGAQFNRPANLKTHTRIHSGKPKYKCEFCG---YKCDRCQASFRYKGNLASHKTVHTGKPYRC 576

QY 208 RSYKQSSLEEH 219
Db 637 TRFRHLQTLKSH 648

RESULT 11
PCT-US94-06669-2
; Sequence 2, Application PC/TUS9406669
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Chaganti, R.S.K.
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
; TITLE OF INVENTION: bcl-6
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/06669
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/074,967
; FILING DATE: 09-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43771-A-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

PCT-US95-09345-13

Query Match 44.3%; Score 1208; DB 3; Length 470;
Best Local Similarity 58.3%; Pred. No. 2.2e-112;
Matches 242; Conservative 50; Mismatches 93; Indels 30; Gaps 11;

QY 111 RTSKMNCDCVGLSCISFNVLVHKSHTGERPPQCNOCGASFTQKGNLLRHKLHGTGE 170
DB 60 RLPNKGLKCDICIGXICIGPNVLVHKSHTGERPPQCNOCGASFTQKGNLLRHKLHGTGE 119
QY 171 KPFKCHLNCYACORDALHTHSHVEKPYKCEFCGRSYKORSLSEHKKRCRAFLONP 230
DB 120 KPFKCHLNCYACRRDALHTHSHVSGPKHCKGCGGRSYKORSLSEHKKRCRCHNYLES 179
QY 231 DL-GDAASV---EARIK-----AEMGSEALVLDRLASNAVKRKSMPQKFIQEKHCF 281
DB 180 GLPGXXXPVIKETXHXEMAEDLCKIGXERSLVLDRLASNAVKRKSMPQKFLGDK-XLS 238
QY 282 DANYNPGWYKENENMQTRMDQAINNAISYLGAEAPRLVQTPAPTSEWVPVIVSY 341
DB 239 DXPYDSA-XYEREXXMXSHVMDXAINNAINYLGAESLRPLVQTPPG-XSEYVVPVISP 296
QY 342 PIALRADPMGAPOMEKKR-----ILLPEKILPSERGLSPNNSAODSTDSDSNHE 393
DB 297 QL---HXSSXCPNSNSAQXAVXXLLLSKAKXXVXSERASPSNQCQDSTDSTESNKE 353
QY 394 DRQH--LYQSHVLPQARNGMPLKEVPRSFELKPPPICULRDSIKVINKEGEYMDVFR 451
DB 354 EQRSGLIYLTNIX--XXAXXXXLKEEXRAYXXLRAASENSQDAXRVVSTSGEOKYVK 411
QY 452 CDHCHVFLDYWFYIHM---CCHGRDPFECNMGCGYRSHDRYEFSSHIARGEHR 503
DB 412 CEHCRVFLDHYMYTHIMXXXGCHGRDPFECNMGCGYRSHDRYEFSSHIARGEHR 466

RESULT 4

US-08-933-750C-4
; Sequence 4, Application US/08933750C
; Patent No. 5932442

GENERAL INFORMATION:

; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933-750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 338 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 9337
; US-08-933-750C-4

Query Match 10.6%; Score 289.5; DB 2; Length 338;
Best Local Similarity 26.7%; Pred. No. 6.6e-21;
Matches 82; Conservative 49; Mismatches 109; Indels 67; Gaps 12;

QY 1 MEDIQTV--ELKSTEQPLPTE-----SPDAL--NDYSLPKPHEIENVDSREAP 46
DB 60 VOEVQDTVLKQMSAQEKDLPQKKHFDNRESQANSGLDITNOVSLQK-----IDNPESQ 113
QY 47 ANEDDAGEDSMKV-----KDEYSRDENI--MKPEPMGDABESEMPSYAREYSDEY 97
DB 114 AN--SGALDTNQVLLHAKIPPRKRLKRDQSVKSMKHSNRVKIHQKSCERQKAKENGCR 170
QY 98 SI--KLERHVPYDNRSTSGKMCNDCVGLSCISFNVLVHKSHTGERPPQCNOCGASFT 155
DB 171 KTFSRSTKQITFIHKGSGVQCRCSECGKIFRNPYFSVHKKIHTGERPYVCQCGKGFV 230
QY 156 QKGNLLRHKLHKTGEKPFKCHLNCYACORRALTGHLRTHSHVEKPYKCEFCGRSYKORSS 215
DB 231 QSSSLTQHVHSGERPEFCECGRTNDRSAISQHLRTHTGCAKPYKQDCGCKAFRQSSH 290
QY 216 LEEHK-----ERC-RAFLQNPDLGDAASVEARHKAEMGSEALVLDRLASNVA 263
DB 291 LIRHQRTHTGERPYACNKGKAKFTQSSHL-----IGHQR-----THNRT 329
QY 264 KRKSSMP 270
DB 330 KRKKKQP 336

RESULT 5

US-08-933-750C-17
; Sequence 17, Application US/08933750C
; Patent No. 5932442

GENERAL INFORMATION:

; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

QY 366 PEKILPSERGLSPNNSQDSTDDTDSDNEDRQH--LYQOSHVVLPQARNGMPLLKVEPRSF 423
Db 424 KAKSVSSREASPSNCSODSTDTESNAEORSGLIYLTNH-INPHARNGL-ALKEEQRAY 481
QY 424 ELLKPPPTCLRLDSIKVNKEGEVMDVFCRCHVFLFDYVNFTHM--GCHGFRDPFEC 480
Db 482 EVLRAASENSODAFRVYSTSGEQLKVYKCEHCRVFLFDHVMYTHMCHGCHGFRDPFEC 541
QY 481 NMCYGRSHDRYEFSSHIARGHR 503
Db 542 NMCYHSDQDRYEFSSHITRGEHR 564

RESULT 2
US-08-465-590-153
; Sequence 153, Application US/08465590
; Patent No. 5824770
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,590
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal
US-08-465-590-153

Query Match 44.3%; Score 1208; DB 2; Length 470;
Best Local Similarity 58.3%; Pred. No. 2.2e-112;
Matches 242; Conservative 50; Mismatches 93; Indels 30; Gaps 11;

QY 111 RPTSGKMNCDVGLSCISFNVLWVHKRSHGTGERPFQCNQCGASTQGNLLRHKLTGE 170
Db 60 RLPNGKLKCDTCGXCIGPNVLWVHKRSHGTGERPFQCNQCGASTQGNLLRHKLSGE 119
QY 171 KPFKCHLCNACQRRDALTGHLRTHSVKPYKCEFCGRSKYKORSSLEEHKRCRAFLQNP 230

Db 120 KPFKCHLCNACQRRDALTGHLRTHSVGPHKCGYCGRSYKORXSLEEHKRCRCHNYLES 179
QY 231 DL-GDAASY---EARHIK-----AEMGSEALVLDRLASNAVKRKSMPQKFTGEKRHC 281
Db 180 GLPGXXXPKVKEETXHXEMAEDLCKIGXERSLVLDRLASNAVKRKSMPQKFTGDK-XLS 238
QY 282 DANYNPGYMYKRENNEMQTRMDQAINNAISYLGAFAFRPLVOTPPAPTSEMVPTSSVY 341
Db 239 DXPYDSA-XYEKEXMMXSHVMDXAINNAIYLGAESLRPLVOTPPG-XSEVVPVISPMY 296
QY 342 PIALTRADMPGAPQEMEKRR-----ILLPEKILPSERGLSPNNSQDSTDDTDSDNHE 393
Db 297 QL---HXXSXGXPRSNHSAQDXYVXXLLLSLAKAXVXSERASPSNCSODSTDTESN 353
QY 394 DRQH--LYQOSHVVLPQARNGMPLLKVEPRSFELLKPPPICLRDSIKVINKEGEVMDVFR 451
Db 354 EORSGLIYLTNHIX--XXAXXXXLKEEXRAYXXLRAASENSQDAXRVYSTSGEQXKVYK 411
QY 452 CDHCHVFLFDYVNFTHM--GCHGFRDPFECNMGYRSHDRYEFSSHIARGHR 503
Db 412 CECRVFLFDHVMYTHMXXXGCHGFRDPFECNMGYHSDQDRYEFSSHITRGEHR 466

RESULT 3
PCT-US95-09345-13
; Sequence 13, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MGP-027PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal

PS Disclosure; Page 75-76; 102pp; English.
CC The sequence of an Ikaros protein (R92021) is provided in the
CC specification. Ikaros protein is a master regulator of
CC hematopoietic differentiation and a major determinant in lymphocyte
CC differentiation. Isoforms of Ikaros (see R92014-19) arise from
CC differential splicing of Ikaros gene transcripts. Transgenic animals,
CC pref. mice, having a mutated Ikaros transgene, esp. a mutation that
CC alters the DNA binding domain of the Ikaros protein, are used as
CC models to determine the effects of treatments for immune or nervous
CC system disorders.
SQ Sequence 470 AA;

Query Match 44.3%; Score 1208; DB 1; Length 470;
Best Local Similarity 58.3%; Pred. No. 2.5e-100;
Matches 242; Conservative 50; Mismatches 93; Indels 30; Gaps 11;

QY 111 RPTSGKMNCDVGLGICISFNVLVHVKHSHTGERPFQCNQCGASFTQKGNLRIKHLHTGE 170
DB 60 RLPNGKLKDCICGIXCIGPNVLVHVKHSHTGERPFQCNQCGASFTQKGNLRIKHLHSGE 119
QY 171 KPFKCHLCNACORRDALDGLRTHSVKPYKCEFCGRSYKQSSLEEHKERCRAFLQNP 230
DB 120 KPFKCHLCNACRRDALDGLRTHSVKPYKCEFCGRSYKQSSLEEHKERCCHNYLES 179
QY 231 DL-GDAASV---EARHIK-----AEMGSEALVLDRLASNVAKKSSMPKFFGKXKHC 281
DB 180 GLPGXXXPVTKETXHXEMAEADCKIGXERSVLDRLASNVAKKSSMPKFFGDK-XLS 238
QY 282 DANYNPGYMEKENEMOTRMDDAINAISYLGAERPLVOTPPAPTSEMPVPISSVY 341
DB 239 DXPYDSA-XYEKEXXMXSHVMDXAINAINLYLGAESRLPLVOTPPG-XSEVPVPISPMY 296
QY 342 PIALTRADMPGAPQEMKKR-----ILLPEKILPSEGLSPNNSAQDSTDTSNHE 393
DB 297 QL---HXXSXGPRNSHQAQXAVXLLLSKAKXVXSERASPSNCSQDSTDTSNHE 353
QY 394 DRQH--LYQOSHVVLPQARNGMPLLEKVPFSFELLKPPICLRDSIKVINKEGVMDVFR 451
DB 354 EORSGLIYLTNHIX--XXAXXXLKEEXRAYXXLRAASQNSQDAXRVVSTSGEQXKVK 411
QY 452 CDHCHVLFLDYVMTIRM---GCHGFDPFCNMGYRSHDRYEFSSHARGHR 503
DB 412 CEHCRLVFLDHVMTIHXXXGCHGFDPFCNMGYRSHDRYEFSSHARGHR 466

RESULT 12

W72678
ID W72678 standard; Protein; 470 AA.
AC W72678;
DT 14-JAN-1999 (first entry)
DE Ikaros protein general formula.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; immune system; corpus striatum; AIDS;
KW Alzheimer's disease.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc_difference 1 /note= "any amino acid"
FT Misc_difference 2 /note= "any amino acid"
FT Misc_difference 74 /note= "any amino acid"
FT Misc_difference 163 /note= "any amino acid"
FT Misc_difference 184 /note= "any amino acid"
FT Misc_difference 185 /note= "any amino acid"
FT Misc_difference 186 /note= "any amino acid"
FT Misc_difference 194 /note= "any amino acid"
FT Misc_difference 194 /note= "any amino acid"

FT Misc_difference 196 /note= "any amino acid"
FT Misc_difference 207 /note= "any amino acid"
FT Misc_difference 236 /note= "any amino acid"
FT Misc_difference 240 /note= "any amino acid"
FT Misc_difference 246 /note= "any amino acid"
FT Misc_difference 251 /note= "any amino acid"
FT Misc_difference 252 /note= "any amino acid"
FT Misc_difference 255 /note= "any amino acid"
FT Misc_difference 261 /note= "any amino acid"
FT Misc_difference 285 /note= "any amino acid"
FT Misc_difference 300 /note= "any amino acid"
FT Misc_difference 301 /note= "any amino acid"
FT Misc_difference 302 /note= "any amino acid"
FT Misc_difference 304 /note= "any amino acid"
FT Misc_difference 306 /note= "any amino acid"
FT Misc_difference 316 /note= "any amino acid"
FT Misc_difference 319 /note= "any amino acid"
FT Misc_difference 320 /note= "any amino acid"
FT Misc_difference 329 /note= "any amino acid"
FT Misc_difference 331 /note= "any amino acid"
FT Misc_difference 352 /note= "any amino acid"
FT Misc_difference 367 /note= "any amino acid"
FT Misc_difference 368 /note= "any amino acid"
FT Misc_difference 369 /note= "any amino acid"
FT Misc_difference 371 /note= "any amino acid"
FT Misc_difference 372 /note= "any amino acid"
FT Misc_difference 373 /note= "any amino acid"
FT Misc_difference 374 /note= "any amino acid"
FT Misc_difference 375 /note= "any amino acid"
FT Misc_difference 380 /note= "any amino acid"
FT Misc_difference 384 /note= "any amino acid"
FT Misc_difference 385 /note= "any amino acid"
FT Misc_difference 397 /note= "any amino acid"
FT Misc_difference 407 /note= "any amino acid"
FT Misc_difference 430 /note= "any amino acid"
FT Misc_difference 431 /note= "any amino acid"
FT Misc_difference 432 /note= "any amino acid"

CC This is the amino acid sequence of human Ikaros protein isoform
 CC hIK-1, deduced from a cDNA clone (see V42806) obtained from a
 CC Jurkat T cell line cDNA library. Native Ikaros is active in the
 CC early stages of lymphocyte differentiation, binding to and
 CC activating the CD3-delta gene enhancer (see V42804). Proteins
 CC of the human Ikaros family (see also W0969 and W0971) are
 CC isoforms that arise from differential splicing of Ikaros gene
 CC transcripts, and contain different combinations of zinc fingers.
 CC They are expressed primarily in T cells in the adult and may play a
 CC role as a genetic switch regulating entry into the T cell lineage.
 CC The human and murine sequences (see also W0963 and W0965-68) are
 CC very similar. The invention provides Ikaros nucleic acids, vectors
 CC and host cells expressing Ikaros polypeptides. These can be used
 CC to treat T and B cell diseases (e.g. immune deficiencies caused by
 CC drugs, radiation or cancer), to control expression of heterologous
 CC genes placed under control of an Ikaros-responsive element, to
 CC treat nervous system diseases (e.g. Alzheimer's disease) and to
 CC modulate cell division, amplification or differentiation, especially
 CC in haematopoietic cells. Some Ikaros isoforms are antagonistic of
 CC others and may be used to inhibit interaction with DNA sequences.
 SQ Sequence 461 AA;

Query Match 46.9%; Score 1279.5; DB 1; Length 461;
 Best Local Similarity 61.5%; Pred. No. 9.4e-107; Indels 31; Gaps 12;
 Matches 254; Conservative 59; Mismatches 69;

QY 111 RPTSGKMNCDVCGLSICISFNVLWVHKRSHTERPFQCCGASFTQKGNLLRHKLHTE 170
 DB 56 RLPNGKLCDCIGICIGPNVLWVHKRSHTERPFQCCGASFTQKGNLLRHKLHTE 115
 QY 171 KPFKCHLCNYACORDALTLGHLRTHSVKPKYKCFECGSRYSKORSLEBHKRCRAFLQNP 230
 DB 116 KPFKCHLCNYACRRDALTLGHLRTHSVKPKYKCFECGSRYSKORSLEBHKRCRAFLQNP 175
 QY 231 DL-GDAASV---EARHIK---AEMGSRALVLDRLASNVAKRKSSMPQKFIKRCRHF 281
 DB 176 GLPTGLTPVKEETKSHMAEDLCKIGERSLVLDRLASNVAKRKSSMPQKFIKRCRHF 234
 QY 282 DANYNPGYMEKENEMMOTRMMDQAINNAISYLGAFAFRPLVQTPTSEMVPVSISSY 341
 DB 235 DTPYDSA-TYEKENEMKSHVMDQAINNAISYLGAFAFRPLVQTPTSEMVPVSISSY 292
 QY 342 PIALTRADMPGARQE-----WEKRIILLPEKILPSEGLSPNNSAQDSTDTDSNHE 393
 DB 293 QLH-RRSE---GTPRSNHSQDSAVEYLLLSKAKLVPSEREASPSNSCQDSTDTESNNE 348
 QY 394 DRQH--LYQQSHVLPQARNMPL-LKEVPRSFELLKPPICLRDSIKVINKEGEVMDVF 450
 DB 349 EQRSLIYLTWHI-----ARRAQRVSLKEHRAVDLLRAASNSQDLRVVSTSGEQMKVY 404
 QY 451 RCDHCHVLFVDYVMTIHMCGHGRDPECNMCGYRSHDRYEFSSHIARGHR 503
 DB 405 KCEHCRVLFVDHVMYTIHMCGHGRDPECNMCGYRSHDRYEFSSHIARGHR 457

RESULT 7
 R46965 ID R46965 standard; Protein; 568 AA.
 AC R46965;
 DT 21-OCT-1994 (first entry)
 DE Ikaros zinc finger protein isoform IK-1.
 KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
 KW corpus striatum; regulatory gene.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT region 1..53
 FT /label= Exons 1/2.
 FT region 54..141
 FT /label= Exon 3.
 FT region 142..247
 FT /label= Exon 4.
 FT region 248..288

FT region /label= Exon 5.
 FT 289..333
 FT /label= Exon 6.
 FT 334..568
 FT /label= Exon 7.
 PN W09406814-A.
 PD 31-MAR-1994.
 PF 14-SEP-1993; U08743.
 PR 14-SEP-1992; US-946233.
 PA (GENO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 94-118387/14.
 PT T-cell pathway regulatory gene, Ikaros - encodes family of unique
 PT zinc finger proteins, useful for treating immune system disorders
 PS Claim 14; Figure 4; 112pp; English.
 CC The Ikaros gene encodes a zinc finger protein which can be used in a
 CC therapeutic composition to treat animals with an immune system
 CC disorder. It may also be used for assessing whether a subject is at
 CC risk for an immune disorder. It is of particular use in treating a
 CC disorder of the corpus striatum.
 SQ Sequence 568 AA;

Query Match 46.9%; Score 1279; DB 1; Length 568;
 Best Local Similarity 49.4%; Pred. No. 1.4e-106; Indels 82; Gaps 18;
 Matches 278; Conservative 72; Mismatches 131;

QY 12 STERQPLPTSPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVKDEYSDRDENI 71
 DB 13 SGRKSPVSDTPDE-GDEPMVPEDLTSTSG--AQQKSDRGNAS-NVKVETQSDENG 68
 QY 72 MKPEPMG-----DAESEMPSYAREYSDYES----- 99
 DB 69 RACEMNGEECAEDLRMLDASGEKMGSHRDQGSALSQGVGIRLPNGKLCDCIGIVCIG 128
 QY 99 --IKLERHVPYDNRSP-----TSCMNCDCVGLSCISFNVLWVHKRS 138
 DB 129 PNVLWVHKRSHTERPFQCCGASFTQKGNLLRHKLHTEGSRALVLDRLASNVAKRKSSMPQKFIKRCRHF 188
 QY 139 HTGERPFQCCGASFTQKGNLLRHKLHTEGSRALVLDRLASNVAKRKSSMPQKFIKRCRHF 198
 DB 189 HTGERPFQCCGASFTQKGNLLRHKLHTEGSRALVLDRLASNVAKRKSSMPQKFIKRCRHF 248
 QY 199 KPYKCFECGSRYSKORSLEBHKRCRAFLQNP---PDLGDAASVEARHIK-----AEMGSE 250
 DB 249 KPHKCGYCGSRYSKORSLEBHKRCRHNLYBSMGLPGYPVPIKEETNHNEMAEADLCKIGAE 308
 QY 251 RALVLDRLASNVAKRKSSMPQKFIKRCRHF--DANYNPGYMEKENEMMOTRMMDQAINN 309
 DB 309 RSLVLDRLASNVAKRKSSMPQKFIKRCRHF--CLSDMPYDSA-NYEKE-DMMTSHVMDQAINN 364
 QY 310 AISYLGAFAFRPLVQTPTSEMVPVSISSYPIALTRADMP-----MGAPOEMEKKRILL 365
 DB 365 AINYLGAESLRPLVQTPTSG--SSEVPVSISSYQLHKKPPSDGPPRSNHSQAQDAVDNLLLS 423
 QY 366 PEKILPSEGLSPNNSAQDSTDTDSNHEDRQH--LYQQSHVLPQARNMPLLKEVPRSF 423
 DB 424 KAKSVSSEREASPSNSCQDSTDTESNAEQRSLIYLTNH-INPHARNGI-ALKEEQRAY 481
 QY 424 ELLKPPICLRDSIKVINKEGEVMDVFCRCHVLFVDYVMTIHM---GCHGRDPEFC 480
 DB 482 EVLRAASNSODAFRVVSTSGEQLKVKYKCHRCRVLFVDHVMYTIHMCGHGRDPEFC 541
 QY 481 NMCGYRSHDRYEFSSHIARGHR 503
 DB 542 NMCGYHSQDRYEFSSHIARGHR 564
 RESULT 8
 R92015 ID R92015 standard; Protein; 461 AA.
 AC R92015;
 DT 09-MAY-1996 (first entry)

QY 407 POAENGMLLKEVPRSEFLLKPPICLRDSIKVINKEGEVMDVFCRDCHVFLDYVMT 466
 DB 416 PHARNGI-ALKEORAYEVLRASSENSQDAFRVSTSGEQLKYKCEHCRVFLDHYMT 474
 QY 467 IHM----GCHGFDPFCNMGCGYRSHDRYEFSSHIARGEHR 503
 DB 475 IHMGCHGCHGFDPFCNMGCGYHSQDRYEFSSHITRGEHR 514

RESULT 5

W70966 ID W70966 standard; Protein; 518 AA.
 AC W70966; DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform m1k-1.
 KW Ikaros; m1k-1; transcription factor; mouse; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT Region 119..139 /note="zinc finger motif"
 FT Region 147..167 /note="zinc finger motif"
 FT Region 175..195 /note="zinc finger motif"
 FT Region 203..224 /note="zinc finger motif"
 FT Region 460..480 /note="zinc finger motif"
 FT Region 491..513 /note="zinc finger motif"
 FT CA2194256-A.
 PN 03-MAR-1998.
 PD 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR N-PSDB; V42808.
 PT New nucleic acid encoding Ikaros protein involved in early differentiation of lymphocytes - existing in several isoforms, and related products, used to treat e.g. immune diseases or cancer and to control cell differentiation
 PS Claim 7; Page 75-77; 158pp; English.
 CC This is the amino acid sequence of murine Ikaros protein isoform m1k-1, deduced from a cDNA clone (see V42808) obtained from a mature murine T cell line E14 library. Native Ikaros is active in the early stages of lymphocyte differentiation, binding to and activating the CD3-delta gene enhancer (see V42804). Proteins of the murine Ikaros family (see also W70963 and W70965-68) are isoforms that arise from differential splicing of Ikaros gene transcripts, and contain different combinations of zinc fingers. They are expressed primarily in T cells in the adult and may play a role as a genetic switch regulating entry into the T cell lineage. The murine and human sequences (see W70964, W70969 and W70971) are very similar. The invention provides Ikaros nucleic acids, vectors and host cells expressing Ikaros polypeptides. These can be used to treat T and B cell diseases (e.g. immune deficiencies caused by drugs, radiation or cancer), to control expression of heterologous genes placed under control of an Ikaros-responsive element, to treat nervous system diseases (e.g. Alzheimer's disease) and to modulate cell division, amplification or differentiation, especially in haematopoietic cells. Some Ikaros isoforms are antagonistic of CC others and may be used to inhibit interaction with DNA sequences.
 SQ Sequence 518 AA;

Query Match 47.1%; Score 1284; DB 1; Length 518;
 Best Local Similarity 53.3%; Pred. No. 4.4e-107;
 Matches 277; Conservative 68; Mismatches 129; Indels 46; Gaps 17;

QY 12 STEEQPLPTESDALNDYSLPKPHEIENVDSREAPANEDEDAGEDSMKVKDEYSDREMI 71
 DB 13 SGKESPVSDTIDE-GDEMPVPEDLSTTSG--AQONSXSDRGMAS-NYKVTQSDSENG 68
 QY 72 MKPEPMG-----DAESEMPTYSYAREYSYKLERHVPYDNSRPTSGKMNCDV 121
 DB 69 RACEMNGEACEDRLMDASGEKMGSHRDQSSALS-----GVGGIRLPNGKLKCDI 121
 QY 122 CGLSCISFNVLVHKRSHTGERPFCNQCQASFTQKGNLLRHIKLTGKPKKCHLCNYA 181
 DB 122 CGIVCIGPNVLMVHKRSHTGERPFCNQCQASFTQKGNLLRHIKLTGKPKKCHLCNYA 181
 QY 182 CORRDALTCHLTHSVKPKYKCEFCGRSYKQSSLEEKRCRAFLQNPDLGDAASV--- 239
 DB 182 CRRRDALTCHLTHSVGKPKKCGYGRSYKQSSLEEKRCRCHNLTSMGLPGVCPVIKE 241
 QY 239 EARHIK-----AEMGSEALVLDRLASNVAKRKSSMPQKFIGEKHCF--DANYNPGMYE 292
 DB 242 ETNHNEMAEEDICKIGAERSLVLDRLASNVAKRKSSMPQKFIGDK--CLSDMPYDSA-NYE 298
 QY 293 KENEMQTRMDOAINNALSILGAERPLVQTPAPTSEMPVVISSVPIALTRADMP- 352
 DB 299 KE-DMTSHVMDQAINNALNYLGAESRLPLVQTPPG--SSEVPVVISSMYQLHKPPSDGPP 356
 QY 352 ---MGAPOQMEKKRILLPEKILPSERGLSPNNSAODSTDTSNHHEDRQH--LYQOSHVVYL 406
 DB 357 RSNHSAQDAVDNLLLSKAKSVSSEREASPSQSDSTDESNAEEQRGLIYLTNH--IN 415
 QY 407 POARNGMLLKEVPRSEFLLKPPICLRDSIKVINKEGEVMDVFCRDCHVFLDYVMT 466
 DB 416 PHARNGI-ALKEORAYEVLRASSENSQDAFRVSTSGEQLKYKCEHCRVFLDHYMT 474
 QY 467 IHM----GCHGFDPFCNMGCGYRSHDRYEFSSHIARGEHR 503
 DB 475 IHMGCHGCHGFDPFCNMGCGYHSQDRYEFSSHITRGEHR 514

RESULT 6
 W70964 ID W70964 standard; Protein; 461 AA.
 AC W70964; DT 11-JAN-1999 (first entry)
 DE Human Ikaros isoform h1k-1.
 KW Ikaros; h1k-1; transcription factor; human; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 64..84 /note="zinc finger motif"
 FT Region 92..112 /note="zinc finger motif"
 FT Region 120..140 /note="zinc finger motif"
 FT Region 148..169 /note="zinc finger motif"
 FT Region 406..426 /note="zinc finger motif"
 FT Region 434..456 /note="zinc finger motif"
 FT CA2194256-A.
 PN 05-MAR-1998.
 PD 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR N-PSDB; V42806.
 PT New nucleic acid encoding Ikaros protein involved in early differentiation of lymphocytes - existing in several isoforms, and related products, used to treat e.g. immune diseases or cancer and to control cell differentiation
 PS Claim 7; Page 70-72; 158pp; English.

PN W09604372-A1.
 PD 15-FEB-1996.
 PF 28-JUL-1995; U09345.
 PR 29-JUL-1994; US-283300.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 96-129389/13.
 DR N-PSDB; T016062.
 PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine effects of treatment for immune and nervous system disorders
 PT Disclosure: Fig 4: 102pp; English.
 CC The sequence of 57.5 kba mouse Ikaros protein mIK-1 (R92017) was deduced from mouse Ikaros cDNA (T16062) isolated from a mature T-cell line E15 library. Ikaros protein is a master regulator of hematopoietic differentiation and a major determinant in lymphocyte differentiation. Other isoforms of Ikaros (see R92014, R92016 and R92018-19) arise from differential splicing of Ikaros gene transcripts. Transgenic animals, esp. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.
 CC Sequence 518 AA;
 SQ

Query Match 47.1%; Score 1284; DB 1; Length 518;
 Best Local Similarity 53.3%; Pred. No. 4.4e-107;
 Matches 277; Conservative 68; Mismatches 129; Indels 46; Gaps 17;

QY 12 STEEPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVDEYSDRDENI 71
 DB 13 SKGESPPVSDTPDE-GDEPMPVPEDLSTSG--AQNSKSDRGMAS-NVKVETOSDEENG 68
 QY 72 MKPEPMG-----DAEESEMPYSYAREYSDYESIKLERHVPYDNRPTSGKMNCDV 121
 DB 69 RACENMGEECAEDLRMLDASGEKMGSHRDOGSALS-----GVGIRLPNGKLKCDI 121
 QY 122 CGLSCISFNVLVHVKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGEKPFKCHLCNYA 181
 DB 122 CGIVCIGPNVLVHVKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGEKPFKCHLCNYA 181
 QY 182 CORRDLTGLHRTSHVSKPYKCEFCGRSYKORSLSLEHKEKRCRAFLONPLDGLDAASV--- 239
 DB 182 CRRDLTGLHRTSHVSKPYKCEFCGRSYKORSLSLEHKEKRCRAFLONPLDGLDAASV--- 239
 QY 239 EARHIK-----AEMGSEALVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYMYE 292
 DB 242 ETNNHMAEDLCKIGAERSLVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYMYE 292
 QY 293 KENEMQTRMDQAINNAISYLGAERPLVQTPPTSEMPVVISVYPIALTRADMP- 352
 DB 299 KE-DNMTSHVMDQAINNAISYLGAERPLVQTPPTSEMPVVISVYPIALTRADMP- 352
 QY 352 ---MGAPOEMEKRRILLPEKILPSEKSLSPNNSAODSTDSDSNHEDROH--LYQOSHVV 406
 DB 357 RSNHSAQDAVDNLLLSKAKSVSSERASPSNSQDSTDSDSNHEDROH--LYQOSHVV 406
 QY 407 POARNMPLLEKRRILLPEKILPSEKSLSPNNSAODSTDSDSNHEDROH--LYQOSHVV 406
 DB 416 PHARNGL-ALKEEQRAYEVLRAAENSODAFRVVTSQGLKVKYKCEHCRVFLDHYVMT 474
 QY 467 IHM---GCHGRFDPPECNMCGYRSHDRYEFSSHIARGHEH 503
 DB 475 IHMGCHGCHGRFDPPECNMCGYRSHDRYEFSSHIARGHEH 514

RESULT 4
 W72674
 ID W72674 standard; Protein; 518 AA.
 AC W72674;
 DT 14-JAN-1999 (first entry)
 DE Mouse Ikaros mIK-1.
 KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;

KW differentiation marker; immune system; corpus striatum; AIDS;
 KW Alzheimer's disease.
 OS Mus sp.
 PN US5824770-A.
 PD 20-OCT-1998.
 PF 05-JUN-1995; 465590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-582821/49.
 DR N-PSDB; V66971.
 PT Ikaros poly-peptide(s) - useful for treating disorders of immune system or corpus striatum
 PS Claim 1: Column 61-66; 111pp; English.
 CC The present invention describes a purified peptide having at least one of the following properties: (a) it stimulates transcription of a DNA sequence under the control of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (c) it competitively inhibits the binding of a naturally occurring Ikaros isoform to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (d) it competitively inhibits Ikaros binding to Ikaros responsive elements; or (e) it inhibits protein-protein interactions of transcriptional complexes formed with naturally occurring Ikaros isoforms. The proteins, provided that they stimulate gene transcription under the control of delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit binding of naturally occurring Ikaros isoforms to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit Ikaros binding to Ikaros-responsive elements and/or inhibit protein-protein interactions of transcriptional complexes with naturally occurring Ikaros isoforms, can be used to treat immune system disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g. Alzheimer's disease. The present sequence represents a specifically claimed mouse Ikaros protein.
 CC Sequence 518 AA;
 SQ

Query Match 47.1%; Score 1284; DB 1; Length 518;
 Best Local Similarity 53.3%; Pred. No. 4.4e-107;
 Matches 277; Conservative 68; Mismatches 129; Indels 46; Gaps 17;

QY 12 STEEPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVDEYSDRDENI 71
 DB 13 SKGESPPVSDTPDE-GDEPMPVPEDLSTSG--AQNSKSDRGMAS-NVKVETOSDEENG 68
 QY 72 MKPEPMG-----DAEESEMPYSYAREYSDYESIKLERHVPYDNRPTSGKMNCDV 121
 DB 69 RACENMGEECAEDLRMLDASGEKMGSHRDOGSALS-----GVGIRLPNGKLKCDI 121
 QY 122 CGLSCISFNVLVHVKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGEKPFKCHLCNYA 181
 DB 122 CGIVCIGPNVLVHVKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGEKPFKCHLCNYA 181
 QY 182 CORRDLTGLHRTSHVSKPYKCEFCGRSYKORSLSLEHKEKRCRAFLONPLDGLDAASV--- 239
 DB 182 CRRDLTGLHRTSHVSKPYKCEFCGRSYKORSLSLEHKEKRCRAFLONPLDGLDAASV--- 239
 QY 239 EARHIK-----AEMGSEALVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYMYE 292
 DB 242 ETNNHMAEDLCKIGAERSLVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYMYE 292
 QY 293 KENEMQTRMDQAINNAISYLGAERPLVQTPPTSEMPVVISVYPIALTRADMP- 352
 DB 299 KE-DNMTSHVMDQAINNAISYLGAERPLVQTPPTSEMPVVISVYPIALTRADMP- 352
 QY 352 ---MGAPOEMEKRRILLPEKILPSEKSLSPNNSAODSTDSDSNHEDROH--LYQOSHVV 406
 DB 357 RSNHSAQDAVDNLLLSKAKSVSSERASPSNSQDSTDSDSNHEDROH--LYQOSHVV 406
 QY 407 POARNMPLLEKRRILLPEKILPSEKSLSPNNSAODSTDSDSNHEDROH--LYQOSHVV 406
 DB 416 PHARNGL-ALKEEQRAYEVLRAAENSODAFRVVTSQGLKVKYKCEHCRVFLDHYVMT 474
 QY 467 IHM---GCHGRFDPPECNMCGYRSHDRYEFSSHIARGHEH 503
 DB 475 IHMGCHGCHGRFDPPECNMCGYRSHDRYEFSSHIARGHEH 514


```

QY 301 RMDQAINNAISYLGAFAEPLVOTPPAPTSEMVPVIVSSVPIALTRADMPGAPQEMEK 360
Db 301 RMDQAINNAISYLGAFAEPLVOTPPAPTSEMVPVIVSSVPIALTRADMPGAPQEMEK 360
QY 361 KRILLPEKILPSERGLSPNNSAODSTDSDNHNEDRHLVQOQSHVVLPOARNGMPLLEKVP 420
Db 361 KRILLPEKILPSERGLSPNNSAODSTDSDNHNEDRHLVQOQSHVVLPOARNGMPLLEKVP 420
QY 421 RSELLKPPICLRDSIKVINKGEVMDVFRCDHCHVFLDYVNFTHMCHGFRDPFEC 480
Db 421 RSELLKPPICLRDSIKVINKGEVMDVFRCDHCHVFLDYVNFTHMCHGFRDPFEC 480
QY 481 NMGYRSHDRYEFSSHTARGEHRAMLK 507
Db 481 NMGYRSHDRYEFSSHTARGEHRAMLK 507

RESULT 2
ID W70971 standard; Protein; 516 AA.
AC W70971;
DT 11-JAN-1999 (first entry)
DE Human Ikaros isoform h1k-1.
KW Ikaros; h1k-1; transcription factor; human; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
KW Alzheimer's disease; therapy; diagnosis.
OS Homo sapiens.
FH Key
FT Region
FT 119..139
FT /note= "zinc finger motif"
FT 147..167
FT /note= "zinc finger motif"
FT 175..195
FT /note= "zinc finger motif"
FT 203..224
FT /note= "zinc finger motif"
FT 461..481
FT /note= "zinc finger motif"
FT 489..511
FT /note= "zinc finger motif"
CA2194236-A.
PD 05-MAR-1998.
PR 02-JAN-1997; 194256.
PR 05-SEP-1996; US-711417.
PA (GCHO ) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR N-PSDB; V42840.
PT New nucleic acid encoding Ikaros protein involved in early
PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer and
PT to control cell differentiation
PS Claim 1; Page 127-129; 158pp; English.
CC This is the amino acid sequence of human Ikaros protein isoform
CC h1k-1, deduced from a cDNA clone (see V42840) obtained from a
CC Jurkat T cell line cDNA library. Native Ikaros is active in the
CC early stages of lymphocyte differentiation, binding to and
CC activating the CD3-delta gene enhancer (see V42804). Proteins
CC of the human Ikaros family (see also W70964 and W70969) are
CC isoforms that arise from differential splicing of Ikaros gene
CC transcripts, and contain different combinations of zinc fingers.
CC They are expressed primarily in T cells in the adult and may play a
CC role as a genetic switch regulating entry into the T cell lineage.
CC The human and murine sequences (see also W70963 and W70965-68) are
CC very similar. The invention provides Ikaros nucleic acids, vectors
CC and host cells expressing Ikaros polypeptides. These can be used
CC to treat T and B cell diseases (e.g. immune deficiencies caused by
CC drugs, radiation or cancer) to control expression of heterologous
CC genes placed under control of an Ikaros-responsive element, to
CC treat nervous system diseases (e.g. Alzheimer's disease) and to
CC modulate cell division, amplification or differentiation, especially
CC in hematopoietic cells. Some Ikaros isoforms are antagonistic of
CC others and may be used to inhibit interaction with DNA sequences.

```

```

SQ Sequence 516 AA;
Query Match 47.7%; Score 1302; DB 1; Length 516;
Best Local Similarity 52.3%; Pred. No. 1.1e-108;
Matches 275; Conservative 75; Mismatches 124; Indels 52; Gaps 16;
QY 8 VELKSTEEQPLTESPDALNDYSLPKPHEIENVDSREAPANEDEAGEDSMKVKDYSDR 67
Db 9 MSFSSGKESPPVSDTPDE-GDEPMPITPDLSTTSGQSSKSDRVVASN---VKVETQSD 64
QY 68 DENIMKPEPMG-----DAESEMPSYAREVSDYESIKLERHVPYDINSRPSGKM 117
Db 65 EENGRACEMNGEACEDLRMLDASGEKMGSHRDQSSALS-----GVGGIRLPNGKL 117
QY 118 NCDVCGLSGISFNVLVHKRSHTGERPFQCNQCGASFTQGNLLRHLRIKHTGEKPFKCHL 177
Db 118 KCDICGIIICIGPNVLVHKRSHTGERPFQCNQCGASFTQGNLLRHLRIKHSGEKPFKCHL 177
QY 178 CNYACORRDALTGHLRTHSVKPYKCEFCGRSYKQSSLEEHKERCRAFLQNPDL-GDAA 236
Db 178 CNYACRRRDALTGHLRTHSVGPKHKGCGYKQRTSLEEHKERCCHNYLESGLPGTLY 237
QY 237 SV---EARHIK----AEMGSEALVLDRLASNVAKRKSSMPQKFIGEKRFCDANYNG 288
Db 238 PVTKETKHSEMAEDLCKISERSLVLDRLASNVAKRKSSMPQKFLGDK-GLSDTTPDSA 296
QY 289 YMYKENENMOTRMDOAINNAISYLGAFAEPLVOTPPAPTSEMVPVIVSSVPIALTRA 348
Db 297 -IYEKENENMKSHVMDQAINNAISYLGAESLPLVOTPPG-GSEVVPVISPVLH-RSS 353
QY 349 DMPMGAPQ-----MEKKRILLPEKILPSERGLSPNNSAODSTDSDNHNEDRHL-L 398
Db 354 E--GTPSRNHSQAQDSAVEYLLLSKALVPSEREASPSNSQDSTDSTESNNEEQSGLI 410
QY 399 YQOSHVVLPQARNGMPL-LKEVPSRPELLKPPICLRDSIKVINKGEVMDVFRCDHCHV 457
Db 411 YLTNHI---ARRAQRVSLKEEHRAVDLLRAASENSQDALRVVSTSGEQMKVKYCEHCRV 466
QY 458 LFLDYVMFTIHMCHGFRDPFECNMGYRSHDRYEFSSHTARGEHR 503
Db 467 LFLDHVMYTIHMCHGFRDPFECNMGYRSHDRYEFSSHTARGEHR 512

RESULT 3
R92017
ID R92017 standard; Protein; 518 AA.
AC R92017;
DT 09-MAY-1996 (first entry)
DE Murine Ikaros protein m1k-1.
KW Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
KW immunocomprised; immune system disorder; nervous system disorder;
KW animal model; m1k-1.
OS Mus musculus.
FH Key
FT Domain
FT 119..140
FT /label= F1
FT /note= "zinc finger domain F1"
FT 147..167
FT /label= F2
FT /note= "zinc finger domain F2"
FT 175..195
FT /label= F3
FT /note= "zinc finger domain F3"
FT 203..224
FT /label= F4
FT /note= "zinc finger domain F4"
FT 460..480
FT /label= F5
FT /note= "zinc finger domain F5"
FT 491..513
FT /label= F6
FT /note= "zinc finger domain F6"

```

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 06:53:38 : Search time 21.23 Seconds
(without alignments)
565.655 Million cell updates/sec

Title: US-09-019-348-2
Perfect score: 2728
Sequence: 1 MEDIOPTVEIKSTEEOPPT
HDBYFSSHARGHRAMIK 507

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2724	99.9	507	1	W15574	Mouse Aiolos polypp
2	1302	47.7	516	1	W70971	Human Ikaros isofo
3	1284	47.1	518	1	R92017	Murine Ikaros prot
4	1284	47.1	518	1	W72674	Mouse Ikaros mik-1
5	1284	47.1	518	1	W70966	Mouse Ikaros isofo
6	1279.5	46.9	461	1	W70964	Human Ikaros isofo
7	1279	46.9	568	1	R46965	Ikaros zinc finger
8	1276.5	46.8	461	1	R92015	Human Ikaros prote
9	1276.5	46.8	451	1	W72672	Human Ikaros. Ikar
10	1265.5	46.4	537	1	R46964	Peptide with Ikar
11	1208	44.3	470	1	R92021	Ikaros protein. Tr
12	1208	44.3	470	1	R42678	Ikaros protein gen
13	1205	44.2	470	1	W70970	Ikaros isoform 1 c
14	1146	42.0	431	1	R92014	Murine Ikaros prot
15	1143	41.9	431	1	R46963	Murine Ikaros zinc
16	1143	41.9	431	1	W72671	Murine Ikaros mik-
17	1143	41.9	431	1	W70963	Mouse Ikaros isofo
18	1024.5	37.6	209	1	W15575	Human Aiolos polypp
19	1020.5	37.4	334	1	R92020	Ikaros protein. Tr
20	1020.5	37.4	334	1	W72677	Ikaros isoform. Ik
21	1020.5	37.4	334	1	W70969	Human Ikaros prote
22	999.5	36.6	390	1	R92018	Murine Ikaros prot
23	999.5	36.6	390	1	W72675	Mouse Ikaros mik-4
24	997.5	36.6	390	1	W70967	Mouse Ikaros isofo
25	993	36.4	432	1	R92016	Murine Ikaros prot
26	993	36.4	432	1	W72673	Mouse Ikaros mik-3
27	993	36.4	432	1	W70965	Mouse Ikaros isofo
28	675	24.7	376	1	W70968	Mouse. Ikaros isofo
29	671	24.6	376	1	R92019	Murine Ikaros prot
30	671	24.6	376	1	W72676	Mouse Ikaros mik-5
31	600	22.0	236	1	W72680	Ikaros protein SEQ
32	588.5	21.6	236	1	W72679	Ikaros protein SEQ
33	537.5	19.7	238	1	W72681	Ikaros protein SEQ
34	313	11.5	56	1	W72688	Ikaros protein SEQ
35	313	11.5	56	1	W72689	Ikaros protein SEQ
36	303	11.1	56	1	W72687	Ikaros protein SEQ
37	273.5	10.0	532	1	W52187	Human BRCA1-associ
38	273	10.0	711	1	W37504	Human OTK18. Novel
39	270.5	9.9	561	1	W82404	Human SRE-ZBP anal
40	269	9.9	412	1	R92068	Renal cancer associ
41	265.5	9.7	706	1	Y07068	BCL-6 zinc finger
42	261.5	9.6	803	1	W817456	Myc-binding zinc-f
43	256.5	9.4	181	1	R83490	Zif268-Zif268 zinc

ALIGNMENTS

RESULT	1
WI5574	
ID	WI5574 standard; Protein; 507 AA.
AC	WI5574;
DT	07-JUL-1997 (first entry)
DE	Mouse Aiolos polypeptide.
KW	Aiolos; transcription activator; immune system; T lymphocyte;
KW	B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW	transgenic animal

OS	MUS sp.	
US	Key	Location/Qualifiers
FH	FT	290..344
FT	domain	/label= Activation_domain
FT		/note= "conserved activation domain of Aiolos"

PN WO9714714-A1.
PR 24-MAR-1997

24-APR-1997:
17-OCT-1996; U16774.

PR 18-OCT-1995; US-005529.

PA (GEHQ) GEN HOSPITAL. CORP
14-MAY-1996; US-017646.
PR

PI Georgopoulos K, M

DR WPI; 97-245047,
DR N-DCDR: T50400

DR N-PSDB; T60490
PT Aiolos polypep

PT mammalian immune system, for the treatment of T cell leukaemia(s)

PS Disclosure: Page 76-78; 115pp; English.

CC Mouse Aiolos (W15574) is a homologue of Ikaros whose expression is

CC restricted to lymphoid lineage. Its amino acid sequence was

CC deduced from a cDNA clone (T60490) obtd. from a mouse spleen cDNA

CC library. A human Aiolos partial polypeptide has also been

CC identified (see also W15575). Aiolos forms homodimers that are

CC potent transcriptional activators. It also forms less potent

CC heterodimers with Ikaros. Recombinant Aiolos can be produced in

CC transformed host cells. Aiolos mixexpressing immune system

CC components can be used to reconstitute the immune system of a

CC mammal. The Aiolos protein, coding sequence or cells expressing

CC Aiolos may be used to treat a disorder in an animal, esp. by gene

CC therapy. Such disorders include T-cell leukaemias and lymphomas.

SQ Sequence 507 AA;

Query Match 99.9% Score 2724 DB 1: Length 507:

Best Local Similarity 99.8%; Pred. No. 2.5e-236;

```
Matches 506; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 MEDIQPTVELKSTEEQPLPTESPDALNDYSLPKPHEIENVDSREAPANEDAGEDSMKV 60

db 1 MEDIOPTVELKSTEROPIPTESPDAINDYSTLPKPKHEIENVDSREAPANEDEAGEDSMKY 60

[illegible]

QY 6I ADEISDRUDENIMKPEPMGDAESENPISYAREYSIDYESIKRLERHVPYDNSKPTSGKMNCD 120

Db 61 KDEYSDRDENIMKPEPMGDAEESEMPYSYAREYSDYESIKLERHVPYDNSRPTSGKMNC 120

QY 121 VCGLSGISFNVLMVHKRSHTGERPFCNQCGASFQKGNLLRHIKLTGEKPFKCHLCNY 180

Db 121 VCGLSCTFNVLMVHKRSHTGERPFQCNQCGASTQKGNLLRHKLHTGKPFCHLCNY 180

QY 181 ACORRDALTGHLRTHSVEKPYKCEFCGRSYKQRSSLEEKKERCRAFTJONPDLGDAASVEA 240

Db 181 ACORRDALGHLRTHSVKPYKCECGRDYKORSSLEEKRCRAFLONPDIGDAASVEA 240

241 RHUKAFWCSEPAI VI DPT ASNUAKPKSSMDOKETICEKPHCFDANVNDGSMYVEKENENEMOT 300

[illegible]

DD 24I KHIAEMGSEALVLDRLASNVAKRKSSMPQKF IGERRHCFDANI NPGMYEKENEMMQT 300

THIS PAGE BLANK (USPTO)


```

Qy 700 cagcagaccacacagtggaagatgaactgcgacgtgtgcgggttatctcgtgcattagctt 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 CAITCGACCTCTTAACGGAAACTAAAGTGTGATATCTGTGGATCATTTGCATCGGCC 115

Qy 760 caacgtcttgatgttcataaagcgaagccataacccgcgaaccccttcacagttaataca 819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 CAATGTGCTCATGTTTCAAAAAGAACCCACACTGGAGAACGCCCTTCCAGTGAATCA 55

Qy 820 gtgcggggcattcttactcagaaggttaacctccctccctcattataaac 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 GTGGGGGCTCATTCACCGAAGGGCAACCTGCTCCGGCACAATCAAGC 5

RESULT 14
H83408 355 bp mRNA EST 13-NOV-1995
LOCUS ys90h12.r1 Soares retina N2b5HR Homo sapiens cDNA clone
DEFINITION IMAGE:222119.5', mRNA sequence.
ACCESSION H83408
NID g1062079
VERSION H83408.1 GI:1062079
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 355)
AUTHORS Hillier,L., Clark,M., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisan,E., Waterston,R., Williamson,A., Woldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On Jan 24, 1995 this sequence version replaced.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 257
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1816 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 257.
Location/Qualifiers
1 .355
/organism="Homo sapiens"
/db_xref="GDB:3850880"
/db_xref="taxon:9606"
/map="17; 17p11.2-pl3; 17p12-17p13"
/clone="IMAGE:222119"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCACTCTCAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3'].
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). The retinas were obtained from a 55 year old
Caucasian and total cellular poly(A)+ RNA was extracted 6
hrs after their removal. The retina RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the

```

```

University of Toronto. Library constructed by Bento
Soares and M.Patima Bonaldo.
BASE COUNT 129 a 67 c 88 g 69 t 2 others
ORIGIN

```

```

Query Match 4.3%; Score 85.2; DB 24; Length 355;
Best Local Similarity 60.3%; Pred. No. 9.5e-11;
Matches 141; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

```

```

Qy 764 gtcttgatgttcataaagcgaagccataacccgcgaaccccttcacagttaatacagtcg 823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 GGCTTGAGAAACATCATAGAACACACAGAGGGAGAGCCCTACAAATGTATCAGTGT 95

Qy 824 ggggcatctttactcagaaggttaacctccctccctcattataaacctgcacacggggaa 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GGAAGAACCTTCGGTCAGAAATCACAACCTCAGAGGACATCATAGAAATTCACACAGGGAA 155

Qy 884 aaaccttttaagtgcacctctgcaactacgcattcccaaggagagatgcgtcacggga 943
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 156 AAACCTTATAAATGTAATCATTTGTGGGAAGCTTTTCAGTCAGAAATCAAACTCAGAGTA 215

Qy 944 caccttagcacacattctgtggaagccgtacaagtgtagtcttcgcggaaga 997
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 CATCAGAACTCATCTACTGCGGAGAAACCTATCATGTGAGGAGTGTGAAAA 269

```

```

RESULT 15
AA764246
LOCUS

```

```

DEFINITION AA764246 356 bp mRNA EST 27-JAN-1998
similar to TR:O13100 O13100 IRAROS HOMOLOG ;, mRNA sequence.
ACCESSION AA764246
NID g2811768
VERSION AA764246.1 GI:2811768
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 356)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On May 8, 1995 this sequence version replaced gi:801088.

```

```

FEATURES
source

```

```

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:651352
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 340.
Location/Qualifiers
1 .356
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1225760"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

```

```

FEATURES
source

```


AA867722/c
LOCUS AA867722 332 bp mRNA EST 16-MAR-1998
DEFINITION vx13e05.r1 Soares 2NMT Mus musculus cDNA clone IMAGE:1264352 5' similar to TR:008900 008900 AIOLOS ;, mRNA sequence.
ACCESSION AA867722
NID 92963167
VERSION AA867722.1 GI:2963167
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2150693.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:666904
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 302.
Location/Qualifiers
1 . 332
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/map="1; 763E11; 5; 5q13.2-5q14.1; Chromosome 7; 21q"
/clone="IMAGE:1264352"
/clone.lib="Soares 2NMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT 62 a 72 c 123 g 74 t 1 others
ORIGIN

Query Match 13.6%; Score 269; DB 39; Length 332;
Best Local Similarity 89.8%; Pred. No. 4.9e-55;
Matches 300; Conservative 0; Mismatches 31; Indels 3; Gaps 1;

Qy 1326 cttctccgccccttagctccagactccgctgtccaccctctgagatgtccagtcacata 1385
||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 CTTTCGCCCTTAGTCTCCAGACTCGCGCTGTCCCACTCTCAGATGTCCAGTCATCA 272
||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1386 gcagtggtacccttagcacttactcggcgccgatatcccaatggggcccgaggaga 1445
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GCAGTGTGTACCCCATAGACTTACNTCCGCCGCGGCTACTTCCCAATGGGGCCCGCAGGAGA 212

Qy 1446 tggaaaagaacggatcctctccagagaagatcttgccttgaacgaggtctgtccc 1505
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 TGGAAAGAAACAGGATCTCTCTCCAGAGAAGATCTGCTTCTGAACGAGGTCTGTCCC 152
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1506 ccaataacagtgcacgagactcacagacacccagacagcaccagagatcgcaacatc 1565
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 CCAATACAGTGCACGAGACTCCACAGACACCCAGACAGCAGGAGATCGCCACATC 92
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1566 tctaccagcaagccacgtgtgtctctcccccagggcccgcaatggatgcctcttctgaag 1625
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 91 T--TCCAGAAAGCCAGCTGTCTCTCCCCAGGCCGCAATGGATGCCTCTTCTGAAGG 35
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1626 agtctccgtcttttgaactctcaagccccc 1659
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 AGGTCCTCGCTCTTTGAACCTCCCGAGCCGCCCTC 1

RESULT 8
AA920812/c
LOCUS AA920812 609 bp mRNA EST 20-APR-1998
DEFINITION vx86f12.r1 Soares 2NMT Mus musculus cDNA clone IMAGE:1282127 5' similar to TR:008900 008900 AIOLOS ;, mRNA sequence.
ACCESSION AA920812
NID g3067591
VERSION AA920812.1 GI:3067591
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407236.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:673927
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 423.
Location/Qualifiers
1 . 609
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1282127"
/clone.lib="Soares 2NMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

Qy 1326 cttctccgccccttagctccagactccgctgtccaccctctgagatgtccagtcacata 1385
||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 CTTTCGCCCTTAGTCTCCAGACTCGCGCTGTCCCACTCTCAGATGTCCAGTCATCA 272
||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1386 gcagtggtacccttagcacttactcggcgccgatatcccaatggggcccgaggaga 1445
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GCAGTGTGTACCCCATAGACTTACNTCCGCCGCGGCTACTTCCCAATGGGGCCCGCAGGAGA 212


```

|||||
Db 301 GAATACAGCCACTATGAACGATTAAGTCGAGAGACAGCTGCCCTATGACACACAGA 360
QY 707 ccaacagtgagggaactgcagctgtgctggtgtatctctgctgattgcttcaacgtc 766
|||||
Db 361 CCAACCGGTGGGAGATGAATCGGAGCTGCGGGTTATCTGCTAGCTTCAACGTC 420
QY 767 ttgatgttcataagcgaagccataccggaagcgcgcgccttcagtgtaatacagtcgagg 826
|||||
Db 421 TNGATGTTTATAGCGAAGCCATACCCGGAACCCCGGTCAGTGGTATCATCAGTCCGG 480
QY 827 gcatcttttactcagaaggt 847
|||||
Db 481 GCATTCTTTACTCAGAAAGT 501

RESULT 4
AI509603/C
LOCUS
DEFINITION
AI509603 480 bp mRNA EST 12-MAR-1999
vx13e05.y1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1264352 5'
similar to gb:L03547 Mouse Ikaros DNA binding protein (MOUSE);,
mRNA sequence.
ACCESSION AI509603
NID 94408508
VERSION AI509603.1 GI:4408508
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 480)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Perron,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL On Jun 5, 1998 this sequence version replaced gi:3188293.
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:666904
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 467.
FEATURES
Location/Qualifiers
1..480
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1264352"
/clone_lib="Soares 2NbMT"
/sex="male"
/tissue_type="thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCACCATCTGAAGTCGGAGCGCGCGGTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two

```

```

rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 88 a 112 c 159 g 121 t
ORIGIN
Query Match 23.8%; Score 472; DB 47; Length 480;
Best Local Similarity 99.0%; Pred. No. 7.2e-104;
Matches 475; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1183 tcagaaattcatcgatgagaagcgcactgttcgatgcacactacaatcccggtacat 1242
Db 480 TCAGAAATTCATCGGTGAGAACGGCACTGCTTCGATGCCAACTACATCCCGGCTACAT 421
QY 1243 gtacgagaaggagaacgagatgatgcagaccgggatggaccagccatcaataacgc 1302
Db 420 GTACGAGAAGGAGAACGAGATGATGCAGACCCGGATGATGGACCAAGCCATCAATAACGC 361
QY 1303 catcgactatctaggggtgaagccttcgcgccttagtccagactccgctgctccac 1362
Db 360 CATCACTATCTAGGGGCTGAAGCCCTTCGCCCTTAGTCCAGACTCCGCCCTGCTCCAC 301
QY 1363 ctctgagatggtcccgatcatcagcagtggtgtaccccatagcacttactcgggctgatat 1422
Db 300 CTCTGAGATGTTCCAGTCATCAGCAGTGTGTACCCCATAGCAGCTTACTCGGGCGGATAT 241
QY 1423 gccaatggggggcccgaggagatggaagaaacggatcctcctgccagagaaagatctt 1482
Db 240 GCCCAATGGGGCCCGCAGGAGATGGAAGAAACGGATCCTCTGCCAGAGAAGATCTT 181
QY 1483 gctctgaacagagtgctgtccccaataaagtgccagagactccacagacacgacag 1542
Db 180 GCCTTCTGAACAGAGTCTGTCTCCCAATAACAGTCCCGAGACTCCACAGACACGACAG 121
QY 1543 caaccacagagatgcccaacatctctaccagcaagcagctgtctctcccccagggccg 1602
Db 120 CAACCAAGGAGTATGCCAACATCTCTACCAAGAAAGCAGCTGTGCTCTCCCGGCGCG 61
QY 1603 caatgggtgcctcttctgaaggagtgctctcttttgaactctctcaagcccccctcc 1662
Db 60 CAATGGGATGCTCTTCTGAAGGAGTGCTCTGCTCTTTTGAACCTCTCAAGCCCGCTCC 1
RESULT 5
AA290536 517 bp mRNA EST 14-APR-1997
LOCUS vb16d03.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:749093
DEFINITION 5' similar to gb:L03547 Mouse Ikaros DNA binding protein (MOUSE);,
mRNA sequence.
ACCESSION AA290536
NID 91937175
VERSION AA290536.1 GI:1937175
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 517)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-NCI Mouse EST Project
Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393523.
Contact: Marra M/Mouse EST Project
WashU-NCI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 08:08:27 ; Search time 437.25 Seconds
(without alignments)
8950.272 Million cell updates/sec

Title: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagagcgccacccgctcgg.....gaactcaaacccacctcgag 1984

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

EST:

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	543.4	27.4	546	48	AI549961
2	503.8	25.4	577	39	AA863970
3	484	24.4	501	48	AI550402
4	472	23.8	480	47	AI509603
5	279	14.1	517	31	AA290536
6	279	14.1	401	49	AI644775
7	269	13.6	332	39	AA867722
8	204	10.3	609	40	AA920812
9	161.8	8.2	936	39	AA866707
10	154.6	7.8	217	41	AI040940
11	113.6	5.7	557	45	AI325349
12	87.2	4.4	479	46	AI428749
13	86.8	4.4	339	38	AA808413
14	85.2	4.3	355	24	H83408
15	84	4.2	356	38	AA764246
16	81.4	4.1	389	28	AA119538
17	81.4	4.1	576	31	AA277407
18	80.2	4.0	382	33	AA400326
19	80	4.0	494	39	AA848430
20	79.2	4.0	424	26	W19169
21	79	4.0	509	27	AA027561
22	79	4.0	271	47	AI511308
23	78.8	4.0	662	26	W28371
24	78.4	4.0	685	26	W27601
25	77.4	3.9	550	27	AA014148
26	76.6	3.9	805	49	AI648952
27	76.2	3.8	552	31	AA289328
28	76	3.8	245	20	Z21189
29	76	3.8	577	43	AI179640
30	75.4	3.8	470	21	RI7304
31	75.4	3.8	467	21	R19256
32	75.4	3.8	431	22	R22824
33	74.8	3.8	512	46	AI06535
34	74.4	3.8	394	27	AA003426
35	74.4	3.8	438	33	AA427841
36	74.4	3.8	485	33	AA451439
37	74.2	3.7	505	29	AA144904
38	74.2	3.7	406	54	HS0009139
39	74	3.7	309	24	H73895
40	73.8	3.7	644	22	H05551
41	73.6	3.7	601	34	AA455712
42	73.6	3.7	497	34	AA499468
43	73.6	3.7	484	42	AI124541
44	73.4	3.7	496	40	AA920944
45	73.2	3.7	392	23	H28944

ALIGNMENTS

RESULT	1	AI549961	546 bp	nrna	EST
AI549961	AI549961	AI549961	546 bp	nrna	EST
LOCUS	vx02e08.x1	Soares 2NbMT Mus musculus cDNA clone IMAGE:1263302 3'			
DEFINITION	Similar to TR:008900 008900 AIOLOS ;, mRNA sequence.				
ACCESSION	AI549961				
NID	94482324				
VERSION	AI549961.1	GI:4482324			


```

; MOLECULE TYPE:  CDNA
; FEATURE:
; NAME/KEY:  CDS
; LOCATION:  1..1170
PCT-US95-09345-5

```

Query Match	11.98;	Score 235.2;	DB 5;	Length 1170;
Best Local Similarity	57.16;	Prod. No. 7e-54;		
Mismatches 595;	Conservative 0;	Mismatches 393;	Indels 54;	Gaps 7;
951	atattaactgcacacgggggaaaaaccttttaagtgtcactctgcactgcacacgatgcc	920		
143	AGAGTGATCGAGGCGATGGGTGAACGGCTTTCCAGTGC AACAGTCTGGGGCCCTCTTTA	202		
921	aaagagagatgcctcacgggacaccttaggacacattctgtggagaagccgtacaagt	980		
203	CCAGAAAGGCAACCTCTCTGGGGACATCAAGCTGCACCTCGGGTGAGAAAGCCCTTCAAT	262		
981	gtgagttctgcggaagaagctacaagcagagaacactccctggagagcacaagaacgct	1040		
263	GCCATCTTTCGAACATATGCTCCCGCGAGGGAGCCCTCACCGCCACCTGAGGAGCG	322		
1041	gccagagcttttctcagaacccctgacctggggagcgtgcaagtgtggagcagaacacaca	1100		
323	ACTCCGTCAATTAAGGAAGAAATAACCAACAACGAGATGCC-----AGAAG	367		
1101	tcaaacgcgagatgggaagtgaagagagctctgctctggagacagattagcaagcaatgtgg	1160		
368	ACCTGTGCAGATAGGAGCAGAGAGGTCCCTTGCTCTGGACAGGCTGCCAAGCAATGTGC	427		
1161	ctaagcgaataagctcgatgcctcagaatactcgttgagaagcggcactgtctcgatg	1220		
428	CCAAACGTAAAGAGCTATGCCTCAGAAATTCITGGAGAGCAAGTGCCTGTGACAGATGC	487		
1221	ccaaactcaatcccggtctacatgttacgagaaggagaaacagatgatgcagaccoggatga	1280		
488	CCTATGACATGTC-----AACTATGAGAAGGAGGATATGATGACATCCCACTGA	538		
1281	tggaccagccatcaataacgccatcacctctaggggctgaagccttcgcgccccctag	1340		
539	TGGACACGCCATCAACAATCCATCACTACCTGGGGGCTGAGTCCCTGGCCCATTTGG	598		
1341	tccagatccgcctgtccaccctctgagatggttcccagtcacatcagcagtggtaccacca	1400		
599	TGCAGACACCCCGGT----AGCTCCGAGGTGGTGCCAGTCATCAGCTCCATGTACCAGC	655		
1401	tagcatttactcgggcgatgccaatggg-----gcccgacagagatgg	1448		
656	TGCACAGCCCCCTCAGATGGCCCCCAGCGTCCCAACCATTCACACAGGAGGCCCTCCCGA	715		
1449	aaagaaacggatcctctgcagagaagatcttgcttctgaacgaggtctgtccccca	1508		
716	ATAACTTGCTGCTGTGCCAAGGCCAAGTCTGTGTCATCGGAGGAGGCCCTCCCGA	775		
1509	ataacagtgcacagagatccacagacacccagacagcaacacagagg---atcgcaaatc	1565		
776	GCAACAGCTGCCAAGACTCCACAGATACAGAGAGCAACGCGGAGGAACAGCGACGGCC	835		
1566	tctaccagcaagccaggtggtctctcccccagggcccgcaatgggagtgctctctgaagg	1625		
836	TTATCTACCTAACCAACCATCAACCCGATGCACGCAATGGGGTGGC---TCTCAAGG	892		
1626	aggctccctgcctctttgaaactctcgaagccccctcccatctgctgagggactccatca	1685		
893	AGGAGCAGCGCGCTACGAGGTGTGAGGGGGGCTCAGAGAACTCGCAGGATGCTTCC	952		
1686	aagtgatcaacaagaaggaggtgatggatgttttcgatgtgtacacatgcacagtcgc	1745		
953	GTGTGGTCAGCACAGTGGCAGCAGCTGAAGGTGACAAAGTGCAGAACTGCCCGGTGC	1012		
1746	tcttcttagattatgtgatgttcacattccacatg-----gggtgcacatggtttcc	1796		

Db 1013 TCCTCTGGGATACAGCTCATGTATACATTCACATGGGCTGCATGGCTGCATGGCTTC 107
 Qy 1797 gtatcccttgagtgtaacatgtgtggctatcgcagccacgacgtatgagttctcct 1856
 Db 1073 GGGATCCCTTGAGTGTACATGTGTGGTATCAGAGCCAGGACAGGTACGAGTTCAT 1132
 Qy 1857 ctcacatgcgcagaggagca 1878
 Db 1133 CCCATATCACGGGGGGAGCA 1154
 RESULT 13
 US-08-465-590-4
 ; Sequence 4, Application US/08465590
 ; Patent No. 5824770
 ; GENERAL INFORMATION:
 ; APPLICANT: Georgopoulos, Katia A.
 ; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
 ; NUMBER OF SEQUENCES: 164
 ; CORRESPONDENCE ADDRESS: 164
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 STATE STREET, Suite 510
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII (text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/465,590
 ; FILING DATE: 05-JUN-1995
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/238,212
 ; FILING DATE: 02-MAY-1994
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/121,438
 ; FILING DATE: 14-SEP-1993
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/946,233
 ; FILING DATE: 14-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Paul L.
 ; REGISTRATION NUMBER: 35,695
 ; REFERENCE/DOCKET NUMBER: MPG-006C2DV
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1296 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1296
 ; US-08-465-590-4

[illegible]

```
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1170
US-08-465-590-6

Query Match 11.9%; Score 235.2; DB 3; Length 1170;
Best Local Similarity 57.1%; Pred. No. 7e-54;
Matches 595; Conservative 0; Mismatches 393; Indels 54; Gaps 7;

Qy 861 atattaaactgcacacgaggggaaaaaccttttaagtgcacacctctgcacactacgcgatgcc 920
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 143 AGAGTATCGAGGATGGGTGAACGGCTTTCCAGTGAACCAAGCTCTGGGGCTCCTTTA 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 921 aaagagagatgcgtcacggacacaccttaggacacattctgtgagaaagccgtacaagt 980
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 203 CCCAGAAAGGCAACCTCTCTCGGGACATCAAGCTCAGCTCGGGTGAAGAGCCCTTCAAT 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 981 gtgagttctgcggaagaactcaagcagagagtgctccctggagagcacaaggaacgct 1040
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 263 GCCATCTTTGCAACTATGCTTCCCGCGGAGGAGCGCCCTCACCGGCCACCTGAGGAGCG 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1041 gccagcgttttttcagaacacctgcacctgggggagcgtgcgaagtgtgagcaagacaca 1100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 323 ACTCGCTCATTTAAGGAAGAACTAACCACACAGAGATGCC-----AGAAG 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1101 tcaagcagagatggagtgagagagctctcgttcgacagattagcaagcaatgtgg 1160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 ACTGTGAAGATAGAGCAGAGAGGTCCCTTGTCTGGACAGGCTGGCAAGCAATGCG 427
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1161 ctaagcgaagaagctcgtcctcagaattcctcgtgagagcggcactgctcgtatg 1220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 CCAACGTAAGAGCTCTATGCTCAGAAATTTCTTGGAGACAAGTGCCTGTCAGACATGC 487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1221 ccaactacaatcccgggtacatgtacgagaagagagaagatgatgcagaccocggatga 1280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 CCTATGACAGTGCC-----AACTATGAGAAGGAGGATATGATGACATCCACGTGA 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1281 tggacaaagccatcaataacgcatcagctatctaggggctgaagcccttcggcccttag 1340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 TGGACAGGCCCAACCAACATGCATCACTACCTGGGGGCTGAGTCCCTGGCCCAATTGG 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1341 tcagactcgcctcctccacctctgagatggtcccagtcacacagcagtggtaccacca 1400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 599 TGCACACACCCCGGT---ACCTCGAGGTGGTGCCAGTCAATCAGCTCCATGTACCAGC 655
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1401 tagcacttactcgggcccagatatgccaatgggg-----gccccgcaggagatgg 1448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 TGCACAAGCCCGCTCAGATGCCCCCGCCACGGTCCCAACCATTCAGCAGAGGACGCGGTGG 715
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1449 aaaaagaagagatctctctgcagagaagatcttcccttcgaaagaggtgtgtcccca 1508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 716 ATAATCTGCTGTGTTCCAAAGGCCAAGTCTGTGTATCGGAGCGGAGAGGCCCTCCCGA 775
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1509 aatacagtgccagactccacagacacgcagacgaacacacaggg---atgcgcaacatc 1565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 776 GCAACAGTGCCAAAGACTCCACAGATACAGAGAGGCAACGCGGAGGAAACGCGCAGCGGCC 835
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 12

PCT-US95-09345-5

; Sequence 5, Application PC/TUS9509345

; GENERAL INFORMATION:

; APPLICANT: Georgopoulos, Katia A.

; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 STATE STREET, Suite 510

; CITY: BOSTON

; STATE: MASSACHUSETTS

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII (text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/09345

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/283,300

; FILING DATE: 29-JULY-94

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/238,212

; FILING DATE: 02-MAY-94

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/121,438

; FILING DATE: 14-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/946,233

; FILING DATE: 14-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Paul L.

; REGISTRATION NUMBER: 35,695

; REFERENCE/DOCKET NUMBER: MGP-027PC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1170 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear


```
RESULT 9
US-08-465-590-8
; Sequence 8, Application US/08465590
; Patent No. 5824770
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,590
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1002
US-08-465-590-8

Query Match 13.6%; Score 269.6; DB 3; Length 1004;
Best Local Similarity 58.9%; Pred. No. 4.1e-63;
Matches 602; Conservative 0; Mismatches 369; Indels 51; Gaps 6;

Qy 878 ggggaaaaacccctttaagtgtcacctctgcaactcgtacgcatgcaaaaggagagatgcgtc 937
Db 1 GGAGAACGCCCTTCAGTGCATCAGTCCGGGGCTCATTCACCCAGAAAGGCCAACCTG 60
Qy 938 acgggacaccccttaggacacattctgtggagagccgtacaaagtgtgagttctgcggaaga 997
Db 61 CTCGGCACATCAAGCTGATCCGGGGAGAACCCCTTCAAAATGCCACCTCTGCAACATAC 120
Qy 998 agctcaagcagagagctccctgtgagggagcacaaggaacgtgcgagcttttcttcag 1057
Db 121 GCTGCCCGCGGAGGACGCCCTCACTGGCCACCTGAGGACGCACTCCCGTCATTAAAGAA 180
Qy 1058 aaccctgacctggggagcgtcaagtgtggaggcaagacacatcaaaagccgagatgga 1117
Db 181 GAACTAACACAGTGAATGCG-----AGAACTGTGTGAAGATAGGA 225

1118 agtgagagagctctcgtctctgacagattagcaagcaatgtggctaaagcaaaagctcg 1177
226 TCAGAGAGATCTCGTGTGACAGACTAGCAAGTAATGTGCGCAACAGTAAGAGCTCT 285
1178 atgectcagaaattcattcgtgtgagaagcgcactgcttcgatgccaatacaatccccgc 1237
286 ATGCTTCAGAAATTTCTTTGGGACAAGGGCCTGTCCCGACACGCCCTACGACAGTGC--- 342
1238 tacatgtacgagaagagaacagatgatgagaccggagatgagccagccagccatcaat 1297
342 --CAGGTACGAGAGGAGAACAAATGATGAATGCCAGTGTGAGGACCAAGCCATCAAC 399
1298 aacgccatcagctatctaggggctgaagccttcgccccttagtccagactccgcgtcgt 1357
400 AAGGCATCAACTACTCTGGGGCGAGTCCCTGCGCGCGTGGTGCAGACGCCCGCG-- 458
1358 ccacccctcagatgggtcccsagtcatacagcagtggtgtatcccat-----agca 1405
458 -GCGGTTCCGAGGTGTCTCCCGTATCAGCCCGATGTATCCAGCTGCACAGCGCTCGGAG 516
1406 ctactcgggcccgatatgccaatggggcccgagagatggaaaagaaacgagatcctc 1465
517 GGCACCCCGCGCTCCCAACCACTCGGCCAGACAGCGCGTGGAGTACTCTGCTGCTC 576
1466 ctgccagagaagatcttgcctctgaacgaggtctgtcccccataaacagtgcacagac 1525
577 TCCAAGGCCAAGTTGGTGCCTCGGAGCGGAGCGCTCCCGAGCAACAGCTGCCAAGAC 636
1526 tccacagacacgcagcaaccacagagatc-----gccaaatctctacacgaagc 1579
637 TCCACGGACACCGAGAGCAACACAGGAGGACGCGCGTCTTATCTACTTGACCAAC 696
1580 cagtggtctctcccccagggccgcaatgggagtcctcttgaagaggtccctcgctct 1639
697 CACATCGCCCGACGGCGCAACGGGTCTG-----CTCAGGAGGAGGACCGCGCC 747
1640 ttgaactcctcaagccccctccatctgctgagggagcactccatcaaaagtatcaacaa 1699
748 TAGACCTGCTGCGCGCGCGCTCCGAGAACTCGCAGGACGCGTCCGCGTGTGACAGACC 807
1700 gaaggggaggtgagtggttttcgatgtgacacactgcacgctccttctcctagattat 1759
808 AGCGGGAGCAGATGAAGGTGTACAAGTGCAGAACACTGCCGGGTCTCTCTCGATCAC 867
1760 gtgatgttcacatccacatgggtgcccattgttcctggtatccctttgagtgtaacatg 1819
868 GTCATGTACACCATCCACATGGGTGCGACCGCTTCCGTGATCTCTTTGAGTGCACATG 927
1820 tgtggctatcgaagccacgcatcgctatgagttctcctctcacatcgccagaggagagcac 1879
928 TCGGCTACACAGCAGGACCGGTACGAGTTCTGTCGCACATAACCGGAGGGAGCAC 987
1880 ag 1881
988 CG 989

RESULT 10
PCT-US95-09345-7
; Sequence 7, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
```



```
Qy 1216 cgtatgacactacacccggtacatgtacgagaagagacgagatgatgcagacccg 1275
Db 1089 CATCCCTATGACAGTGCCTAACTATG-----AGAAGGAGGATATGATGACATCCCA 1139
Qy 1276 gatgatgacacacccatcaataacgcatcagctatctaggggtgaagccttccgcc 1335
Db 1140 CGTGATGGACCGCCATCAACATGCATCACTACTCTGGGGCTGAGTCCCTGGCCGC 1199
Qy 1336 cttagtcagactccgctgtccacactctgagatggtccacatcatcagcagtgta 1395
Db 1200 ATTGGTCGACAGCCCCCGGT---AGTCGAGGTGTGTCAGTCATCAGTCCATGTA 1256
Qy 1396 cccatagcacttactcggcgatagtcgaatggggccc-----ccgagga 1443
Db 1257 CCAGCTGCACAAGCCCCCTCAGATGCGCCCGCCACGCTCCCAACCATTCAGCACAGGACGC 1316
Qy 1444 gatgaaagaaacggtatccctctccagagagaatcttgcctctgaacgaggtctgc 1503
Db 1317 CGTGGATAACTTGTGCTGTCTCCAGGCCAAGTCTGTGTATCGGAGCGAGAGGCTC 1376
Qy 1504 ccccaataacagtgccagagactccacagacacccagcagcaaccacaggg---atcgcca 1560
Db 1377 CCCGAGCAACAGCTGCCAAGACTCCACAGATACAGAGAGCAACCGGAGGAACACAGCGCAG 1436
Qy 1561 acatctaccagcaagcagcgtggtctcccccagggccgcaatggatgctcttct 1620
Db 1437 CGGCTTATCTTACCTACCAACCACTCAACCCCGCATGCAGCAATGGGTGGC---TCT 1493
Qy 1621 yaagaggtccctcgtctttttgaactctcctcaagcccccctccatctgcctgagggactc 1680
Db 1494 CAAGGAGGAGCAGCGCCCTACGAGGTCTTGAGGGCGCCCTCAGAGAACTCGCAGGATGC 1553
Qy 1581 catcaagtgatcaacaagaagggggggtgatggtgtgttgcagtgtgacccactgcca 1740
Db 1554 CTTCGGTGTGTCTCAGCAGCAGTGGCGAGCAGCTGAAGGTGTACAAAGTGCAGAACTGCCG 1613
Qy 1741 cgtcctctcagatattgatgtttcaccatccacatg-----gggtgcccagtg 1791
Db 1614 CGTCTCTTCTGTGATCAGGTCATGTATACCATTCATCATGGGTGCGCATGGCTGCCATGG 1673
Qy 1792 ttccctgtagccctttgagtgtaacatgtgtggtatctcgaagccacacgcatgctatgatt 1851
Db 1674 CTTTCGGGATCCCTTTGAGTGTAAATGTGTGTTATCAGACCCAGGACAGGATACGAGTT 1733
Qy 1852 ctctctcactatcgcagagagagagcagacagcagcagcagcagcagcagcagcagcagc 1899
Db 1734 CTCATCCCATATCAGCGGGGGGAGCATCGTTACCACTGTAGCTAAAC 1781
```

```
RESULT 5
PCT-US95-09345-4
; Sequence 4, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
```

```
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MGP-027PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO. 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 223..1776
; PCT-US95-09345-4

Query Match 19.1%; Score 379.6; DB 5; Length 2049;
Best Local Similarity 61.8%; Pred. No. 2.1e-92;
Matches 771; Conservative 0; Mismatches 414; Indels 63; Gaps 8;

Qy 700 cagcagacacacacagcgggaagatgaactgcagcgtgctgcgggttatcctgcattagctt 759
Db 549 CATCGACTTCTTACGGAAAACTAAAGTGTGTATCTGTGGATCGTTTGCATCGGCC 608
Qy 760 caacgtcttgatggttcataaagcgaagccataccgcgcgaacccccgttccagtgtaatca 819
Db 609 CAATGTCTCATGTGTACAAAAAGAGTCACTACTGTGTGAACGGCTTTTCCAGTGCAACCA 668
Qy 820 gtgcggggcatctttactcagaagaagtaacctcctcctcgtcattataaactgcacacggg 879
Db 669 GTCTGGGGCTCTCTTACCAGAAAGCAACCTCTCTCGGCACATCAAGCTGCATCTCGG 728
Qy 880 ggaacacccctttaaagtgtcacctctgcacactacgcatgccaaagagagatgcgtcac 939
Db 729 TGAGAACCCCTTCAATGCCATCTTTGCAACTATGCTGCCCGGAGGAGGCCCTTCAC 788
Qy 940 gggacaccccttagacacattctgtgagaagccgtacaaagtgtgagttctgcggaagaag 999
Db 789 CGCCACCTCAGGACGCACCTCCGTGTGTAAGCCTCACAATGTGGATATTGTGCCCGGAG 848
Qy 1000 ctacaagcagagaagctccctgcggagagacaaaggaacgctgcgcagcttttcttcagaa 1059
Db 849 CTATAACACGCAAGCTCTTTAGAGGAGCATAAAGAGCGATGCCACAACACTACTTGGAAAG 908
Qy 1060 ccttgacctggggagc-----ctcgaagtgtggagcgaagacacatcaag----- 1107
Db 909 CATGGGCTTCCGGGGCTGTGCCAGTCATTAGGAAGAAACTAAACCACACAGATGGC 968
Qy 1107 -----ccgagatgggaagtgcagagagctctcctctgcagattagcaagcaa 1155
Db 969 AGAAGACCTCTGTCAAGATAGGAGCAGAGAGGTCCCTTGTCTCTGGACAGGCTGCAACGCA 1028
Qy 1156 tgtggttaagcgaagaagctcgtgcctcagaattcatcgttgagaagcggcactgctt 1215
Db 1029 TGTGCGCAAAACGTAAGAGCTCTATGCTCAGAAATTTCTTTGGAGACAAAGTCCCTGTGAGA 1088
Qy 1216 cgtgccaactacaatcccggtctacatgtacgagaagagagagagatgatgcagacccg 1275
Db 1089 CATGCCCTATGACAGTGCCTCAACTATG-----AGAAGGAGGATATGATGACATCCCA 1139
```



```
QY 1153 caatgtggttaagcgaagaaagctcgatgcctcgaagaattcatcgtggtgagagcggcactg 1212
Db 642 TAATGTGCGCAAAAGCTTAAGAGCTCTTTCCTCGCTCAAAATTTCTTGGGACAAAGGCGCTGTC 701
QY 1213 ctctgagccaactacatacccggtctacatgtacgagaagaaacagagatgatgcagac 1272
Db 702 CGACAGCCCTAGACAGTGC-----CAGGTACGAGAGGAAACGAAATGATGAAGTC 755
QY 1273 ccgagatgggacaaagccatcaataacgcgcacatcagctatctatctgagggtgtaagccttcg 1332
Db 756 CCACGTGATGACCAAGCCATCAACAACGCCATCAACTACTTGGGGCCGAGTCCCTGCG 815
QY 1333 ccccttagtcagactccgctgctcccaactctgagatggtcccaagtcacagcagtg 1392
Db 816 CCGCGTGTGCAGACGCCCGCG---GCGGTTCGAGGTGTCGCCGTGCATCAGCCCGAT 872
QY 1393 gtaccccata-----gcacttactcggccgcatatgccaatggggcccccga 1440
Db 873 GTACCAGCTGCACAGGCGCTCGGAGGCGCACCCGCGCTCCAACCACTCGGCCCCAGGNACG 932
QY 1441 ggagatggaaagaaacggatcctctgccagagaagatcttgcttctgaaagaggtct 1500
Db 933 CGCGGTGGAGTACTGCTGCTCTCTCAAGGCCAAGTTGCTGCCCTCGGAGCGCGAGGC 992
QY 1501 gtccccaataacagtgccaggactccacagacacacagacagacacacagagatc---- 1557
Db 993 GTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAACAGGAGGACGCG 1052
QY 1557 --gccaacatctaccagcaagcagctggtctctcccccagccgccaatggatgcc 1614
Db 1053 CAGCGGTCTTATCTGCTGACCAACCACTGCGCCGAGCGCGCAAGCGGTGTCG----- 1108
QY 1615 tctctgaaagggtccctcgctctctttgaaactcctcaagcccccctccactctgcctgag 1674
Db 1108 ----CTCAAGGAGGAGCACCGCGCTACGACCTGCTGCGCGCGCGCTCCGAGAACTCGCA 1163
QY 1675 ggactccataaagtgatatacaaaagaggggaggtgatgagtggtttcgatgtgacca 1734
Db 1164 GGAACGCGCTTTCGCGTGTGTCAGCACACGCGGGGAGCAGATGAAGGTGTACAAAGTCCGAACA 1223
QY 1735 ctgcacgctctctctagattatgtatgtttcaccatccacatggggtgcccagtggtt 1794
Db 1224 CTGCGGGTCTCTTCTGTGATCAGCTCATGTACACCATCCACATGGGCTGCCACGGCTT 1283
QY 1795 ccdtgatccctttgagtgtaacatgtgtggtctatcgaaagccacgagatgctatgattctc 1854
Db 1284 CCGTGATCCTTTTGAAGTGCAACATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCTC 1343
QY 1855 ctctcacatgccagaggagagacag 1881
Db 1344 GTCGCACATAACGCGGGGAGCACCG 1370
```

RESULT 4

US-08-465-590-5
Sequence 5, Application US/08465590
Patent No. 5824770

GENERAL INFORMATION:
Applicant: Georgopoulos, Katia A.

Title of Invention: IKAROS: A T CELL PATHWAY REGULATORY GENE
Number of Sequences: 164

Correspondence Address:
Address: LAHIVE & COCKFIELD

City: BOSTON
State: MASSACHUSETTS

Country: USA
Zip: 02109

Computer Readable Form:
Medium Type: Floppy disk

Computer: IBM PC compatible
Operating System: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MPG-006C2DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 223..1776
US-08-465-590-5

Query Match 19.1%; Score 379.6; DB 3; Length 2049;
Best Local Similarity 61.8%; Pred. No. 2.1e-92;
Matches 771; Conservative 0; Mismatches 414; Indels 63; Gaps 8;

```
QY 700 cagcagacaaacccagtggaagatgaactgcagctgctggtggttattcctgaattgctt 759
Db 549 CATTCGACTCTCCTAACGGAAACTAAAGTGTGATATCTGTGGGATCGTTTGATCGGGC 608
QY 760 caactctgtatgttcataagcgaagcattaccggcgaaccccggtccagtgtaaca 819
Db 609 CAAATGCTCATGTGTTACAAAAGAGTCACTACTGTGTGAACGGCCTTTCCAGTGCACAA 668
QY 820 gtgctggggcatctttactcagaaggttaacctcctccctcattataaactgcacacgg 879
Db 669 GTCTGGGGCCTCCTTTACCCAGAAAGCAACCTCTCTGCGGCACATCAAGCTGCACTCGG 728
QY 880 ggaataaaccttttaagtgtcacctctgcaactacgcagatcccaagagagatgcgtcac 939
Db 729 TGAGAAAGCCCTTCAAAATGCCATCTTTTGAACATATGCTGCCCGCGGAGGACGCCCTCAC 788
QY 940 gggacaccttagacacattctgtgagagcgtacagtgatgagttcttcggaagaag 999
Db 789 CGGCCACCTTGAGGACCGACTCCCGTTGGTAAGCCTCAAAATGTGGATATTGTGGCCGGAG 848
QY 1000 ctacaagcagagaagctccctgtgagagcacaaggaacgctgcgagcttttttcagaa 1059
Db 849 CTATAAACAGCGAAGCTCTTTAGAGGAGCATAAAGAGCGATGCCAACACTACTTGGAAAG 908
QY 1060 cctcgacctgggggacg-----ctgcaagtggtggaggcaagacacatcaag----- 1107
Db 909 CATGGGCCCTTCGGGGCGTGTGCCCACTATTAAAGGAAGAAACTTAACCAACAGAGATGGC 968
QY 1107 -----cgaagatgggaagttagagactctcctcctggagacagatagcaagaa 1155
Db 969 AGAAGACCTGTGCAAGATAGGAGCAGAGAGGTCCCTTGCTTGGAGACCAAGCTGCAAGAA 1028
QY 1156 tgtggcgaagcaaaaagctcgtcctcagaaattcatcgttgagaagcggcactgctt 1215
Db 1029 TGTGCGCAAAACGTAAGAGCTCTATGCTTCAGAAATTTCTTGGAGACCAAGTGCCTGTGCA 1088
```


Qy 1050 ccctgacct--gggggacgtgcgaagtgtgaggaagacacatcaaacg----- 1108
Db 744 CATGGCCCTTCGGGACACAGTGTACCCAGTCAATTAAGAAAGAACTAAGCAGTGAAT 803
Qy 1108 -----cgagatgggaagtgagagagctctctgacagattagcaag 1152
Db 804 GGCAGAACACTGTGCAAGATAGATCAGAGAGATCTCTGCTGGACAGACTAGCAAG 863
Qy 1153 caatgtggttaagcgaagaaactgcgtcctcagaataatctcgtgagaagcgcactg 1212
Db 864 TAATGTCCCAACAGTAAGAGCTATGCCCTCAGAAATTTCTGGGCAAGGCCCTGTGTC 923
Qy 1213 cttgagtgccaaactacaatcccggtacatgtacgagaagagacagatgacagac 1272
Db 924 GCACAGCCCTTACGACAGTGC-----CACGTACGAGAAGGAGAACGAAATGATGAAGTC 977
Qy 1273 ccggtatgaggaacaaagccatcaataacgacatcagctatctaggggtgaagccttcg 1332
Db 978 CCACGTGATGACCAAGCCATCAACAGGCCATCACTACCTGGGGCCGAGTCCCTGCG 1037
Qy 1333 ccccttagtcagactccgctgtctccacactctgagatggtccagtcacagcagtgt 1392
Db 1038 CCGCGTGTGTCAGACGCCCGCGG---CGGTTCGAGGTGTCTCCGGTTCATCAGCCCGAT 1094
Qy 1393 gtaccccaata-----gcacttaactcgggcccagatgccaatgggggcccccgca 1440
Db 1095 GTACAGCTGCACAGCGCCTCGGAGGACACCCCGCTCCACCACTCGGCGCCAGGACAG 1154
Qy 1441 ggagatggaagaaagcgtctctctcagagaagatctgtctctcagcagaggtct 1500
Db 1155 CGCGTGTGAGTACCTGTGCTCTCCAAAGGCCAAGTGTGTCTCGGAGCGCGAGGC 1214
Qy 1501 gtcccccataaagtgccaggaagctccacagacacagacagcagcagcagggatc--- 1557
Db 1215 GTCCCGGAGCAACAGCTGCCAAGACTCCACGAGCACCCGAGAGCAACACAGGAGCAGCG 1274
Qy 1557 --gcaacatctctaccagcaaaagcagcgtgtctcccccagcccgccaatggatgcc 1614
Db 1275 CAGCGGTCTTATCTACCTGACCAACCATCGCCCGACCGCGCAACCGGTGTGTCG---- 1330
Qy 1615 tctctgagagaggtccctcgtctcttttgaactctcctcaagcccccctccatctcctgag 1674
Db 1330 ---CTCAAGGAGGAGCACCGCGCTACGACCTCGTGGCGCGCCCTCCGAGAACTCGCA 1385
Qy 1675 ggaactccataaagtatcaacaagaagggaggtgagatgtgttctgagtgacca 1734
Db 1386 GGACGCGTTCGCGGTGTGTCAGCACAGCGGGGAGCAGATGAAGGTGTACAAGTGCAGACA 1445
Qy 1735 ctgcccagctctctctctagattatgtatgttcacacacacatgggtgcccagttt 1794
Db 1446 CTGCGGGTGTCTTCTCTGGATACAGTATGTACACCATCCACATGGGCTGCCACGGCTT 1505
Qy 1795 ccgtgacccctttgagtgtaacatgtgtggtctatcgaagccacagcagcagctatgattctc 1854
Db 1506 CGTGATCTTTTGTAGTGCACCATGTGGGCTACCCACAGCCAGGACCGGTACGAGTCTC 1565
Qy 1855 ctctcaatcgcagagagagagacagagccatgttgaagtgcac 1899
Db 1566 GTGCGACATACCGGAGGAGCAGCCGCTCCACATGACGTAAAGC 1610

RESULT 2

US-08-465-590-3

; Sequence 3, Application US/08465590

; Patent No. 5824770

; GENERAL INFORMATION:

; APPLICANT: Georgopoulos, Katia A.

; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE

; NUMBER OF SEQUENCES: 164

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 STATE STREET, Suite 510

CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MPG-006C2DV
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1386
US-08-465-590-3

Query Match 20.8%; Score 411.8; DB 3; Length 1386;

Best Local Similarity 63.3%; Pred. No. 4.2e-101;

Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

Qy 700 cagcagaccacacagtgaggagatgaactcgacgtgtgctggttattcctgcattagctt 759

Db 162 CATTCGACTTCTTAACGGAAACTAAAGTGTATCTGTGGATCATTTTCATCGGGCC 221

Qy 760 caacgtcttgatgttcaataaggaagccataccggaagcagcccggtccagtgtaataca 819

Db 222 CAATGTGCTCATGTGTTACAAAAGAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCA 281

Qy 820 gtgcggggcactctttactcagaagaagtaacctctcctcgtcatattaaactgcacacgg 879

Db 282 GTCGGGGGCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGTGTGATTCGG 341

Qy 880 ggaataaaccttttaagtgtcacctctgcaactacgcatacccaagagagagatgcctcac 939

Db 342 GGAGAAGCCCTTCAATGTCCACCTCTGCAACTAGCCCTGCCCGGAGGAGGCCCTTCAC 401

Qy 940 ggaacaccttaggacacattctgtgagaagccgtacaaagtgtgagttctgcggaagaag 999

Db 402 TGCCACCTTGAGGACGACCTCCGTTGGTAAACCTCAAAATGTGGATATTGTGGCCGAAG 461

Qy 1000 ctacaagcagagaagctccctcgagagcacaaggaagcagctgcgagcttttcttcagaa 1059

Db 462 CTATAACAGCGAAGCTCTTTTAGAGNAACATAAAGACCGCTGCCACACTACTTGGAAAG 521

Qy 1060 ccttgacct--gggggacgctgcaagtgtgaggaagacacacatcaaacg----- 1108

Db 522 CATGGCCCTTCGGGACACACTGTACCCAGTCAATTAAGAAAGAACTAAGCAGTGAAT 581

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 06:55:19 ; Search time 44.29 Seconds
(without alignments)
4660.693 Million cell updates/sec

File: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagcagcgcacacgcctcgg.....gaactcaaacccacctcgag 1984

Scoring table: IDENTITY_NUC

Searched: 192659 seqs, 52021692 residues

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS9_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413.8	20.9	1611	5	PCT-US93-08743-3
2	411.8	20.8	1386	3	US-08-465-590-3
3	411.8	20.8	1386	5	PCT-US95-09345-2
4	379.6	19.1	2049	3	US-08-465-590-5
5	379.6	19.1	2049	5	PCT-US95-09345-4
6	346.4	17.5	1788	3	US-08-465-590-2
7	346.4	17.5	1788	5	PCT-US93-08743-2
8	346.4	17.5	1788	5	PCT-US95-09345-1
9	269.6	13.6	1004	3	US-08-465-590-8
10	269.6	13.6	1004	5	PCT-US95-09345-7
11	235.2	11.9	1170	3	US-08-465-590-6
12	235.2	11.9	1170	5	PCT-US95-09345-5
13	191.2	9.6	1296	3	US-08-465-590-4
14	191.2	9.6	1128	3	US-08-465-590-7
15	191.2	9.6	1296	5	PCT-US95-09345-3
16	191.2	9.6	1128	5	PCT-US95-09345-6
17	77.4	3.9	2031	4	US-08-933-750C-53
18	68.2	3.4	7218	2	US-08-232-463-14
19	65.8	3.3	1892	4	US-08-933-750C-66
20	63	3.2	2133	3	US-08-820-170A-11
21	63	3.2	3754	3	US-08-820-170A-12
22	60.2	3.0	3720	1	US-08-074-967-1
23	60.2	3.0	3720	4	US-08-553-541B-1
24	60.2	3.0	3720	5	PCT-US94-06669-1
25	58.4	2.9	1309	4	US-08-933-750C-63
26	58.4	2.9	4252	5	PCT-US95-08429-4
27	53.6	2.7	3810	5	PCT-US95-08429-8
28	47.6	2.4	3132	3	US-08-224-482-3
29	46.4	2.3	2850	3	US-08-224-482-7
30	46.2	2.3	2811	3	US-08-040-548-31
31	46.2	2.3	2811	3	US-08-466-344-31
32	46.2	2.3	1440	3	US-08-224-482-5
33	44.4	2.2	267	3	US-08-040-548-14
34	44.4	2.2	3086	3	US-08-040-548-15
35	44.4	2.2	267	3	US-08-466-344-14
36	44.4	2.2	3086	3	US-08-466-344-15
37	44.4	2.2	3068	3	US-08-224-482-1

38 44 2.2 1161 4 US-08-616-857-1 Sequence 1, Appli
39 43.2 2.2 2043 4 US-08-398-590A-39 Sequence 39, Appli
40 43.2 2.2 5648 5 PCT-US96-03940-1 Sequence 1, Appli
41 43.2 2.2 756 5 PCT-US96-03940-2 Sequence 2, Appli
42 43.2 2.2 1407 5 PCT-US96-03940-3 Sequence 3, Appli
43 43.2 2.2 1090 5 PCT-US96-03940-4 Sequence 4, Appli
44 43.2 2.2 928 5 PCT-US96-03940-5 Sequence 5, Appli
45 43.2 2.2 1791 5 PCT-US96-03940-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1
PCT-US93-08743-3
; Sequence 3, Application PC/TUS9308743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08743
; PRIOR APPLICATION DATA:
; FILING DATE: 14-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1611 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1611
PCT-US93-08743-3

Query Match 20.9%; Score 413.8; DB 5; Length 1611;
Best Local Similarity 63.18; Pred. No. 1.3e-101;
Matches 785; Conservative 0; Mismatches 397; Indels 63; Gaps 7;
Qy 700 cagcagcaccacccagtggaagatgaactgcgcagctgtgcgggtttatccttcattagctt 759
Db 384 CATTGCGACTCTCTACCGGAAACATAAGTGATATCTGTGGATCATTTGCGTGGGCC 443
Qy 760 caacgtcttgatggttcataagcgaagccatcacccggaacccggttccagtgtaaca 819
Db 444 CAATGTGCTATGTTTCAAAAAGAACCCACATGAGAGAGGGCCCTTCCAGTGAATCA 503
Qy 820 gtgcggggcactctttactcagaaggttaacctctcctcgtatataaactgcacacgg 879
Db 504 GTGCGGGGCCCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCTCCGG 563
Qy 880 ggaataaccccttttaagtctacacctctgcaactagcgtacgcaaaagagagatgcgtcac 939
Db 564 GGAGAAGCCCTTCAAAATGCCACCTGTGCACTACGCTTGCCCGGAGGAGCGCCCTCAC 623
Qy 940 gggacaccttgaggacacattctgtggagaagccgtacaagtgtagttcttcggaagaag 999
Db 624 TGGCCACCTGAGGAGCGCACTCCGTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAG 683
Qy 1000 ctacaagcagagagctccctggaggagacacaaggaacgctgcggagctttttcttcagaa 1059
Db 684 CTATAAACAGCGGACGCTCTTTAGAGGAACATAAAGAGGCGCTGCCACAACTACTTGGGAAG 743

RESULT 15

ID T16065 standard; cDNA; 1004 BP.
 AC T16065;
 DT 09-MAY-1996 (first entry)
 DE Ikaros cDNA.
 KW Ikaros; transgene; transgenic animal; transgenic mouse;
 KW immunocompromised; immune system disorder; nervous system disorder;
 KW animal model; ss.
 OS Not specified.
 PN WO9604372-A1.
 PD 15-FEB-1996.
 PE 28-JUL-1995; U09345.
 PR 29-JUL-1994; US-283300.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 96-129389/13.
 DR P-PSDB: R92020.
 PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
 PT severely immuno-compromised and can be used as model to determine
 PT effects of treatment for immune and nervous system disorders
 PS Disclosure; Page 71-73; 102pp; English.
 CC A cDNA clone (T16065) coding for an Ikaros protein (R92020) has
 CC been identified. Transgenic animals, pref. mice, having a
 CC mutated Ikaros transgene, esp. a mutation that alters the DNA
 CC binding domain of the Ikaros protein, are used as models to
 CC determine the effects of treatments for immune or nervous system
 CC disorders.
 SQ Sequence 1004 BP; 235 A; 328 C; 286 G; 155 T;

Query Match 13.6%; Score 269.6; DB 1; Length 1004;
 Best Local Similarity 58.9%; Pred. No. 1.9e-58;
 Matches 602; Conservative 0; Mismatches 369; Indels 51; Gaps 6;

QY 878 ggggaaaaacctttaaagtgtcacctctgcaactacgcatcacaagagagatgcctc 937
 Db 1 GGAGAACGGCCCTCCAGTGGCAATCAGTGGCGGGGCCCTCATTCACCCAGAGGGGCAACTG 60
 QY 938 acgggacaccttagacacattctgtggaagcgcgtacaaagtgtgattcttcgggaaga 997
 Db 61 CTCGGGCACATCAAGCTCATTCCTGGGGAGAGGCCCTCAATGCCACCTCTGCAACTAC 120
 QY 998 agctacaagcagagaagctccctgagagagcacaaggaacgctgcgagcttttctcag 1057
 Db 121 GCCTGCCCGCGGAGCGCCCTCACTGGCCACCTGAGGACGCACCTCGTCAATTAAGAA 180
 QY 1058 aaccctgacctggggagcgtgcaagtgtggaggcaagacacatcaaaagcagatggga 1117
 Db 191 GAACCTAAGCACAGTGAATGGC-----AGAAAGACCTGTGCAAGATAGGA 225
 QY 1118 agtgagagagctctcgtcttgacagattagcaagcaatgtggctaagcgaaaaagctcg 1177
 Db 226 TCAGAGAGATCTCTCGTCTGGACAGACTAGCAAGTAATGTCCGCCAACGTAAGAGCTCT 285
 QY 1178 atgctcagaataatcattcgttgagaagcggcaactgtcttcgatgccaactacaatcccggc 1237
 Db 286 ATGGCTCAGAAATTTCTTGGGGACAAGGGCCCTGTCCGACACGCGCTACGACAGTGC--- 342
 QY 1238 tacatgtacagagaggaagacagatgatgagaccggagatgagaccaagccatcaat 1297
 Db 342 --CAGTACGAGAGAGGAAGAAATGATGAAGTCCCACTGGTGGACCAAGCCATCAAC 399
 QY 1298 aacgccatcagctactaggggctgaagccttcgcgccttagtccagactcccgctgct 1357
 Db 400 AACGCCATCAACTACCTGGGGCCCGAGTCCCTGGCCCGCTGGTGACAGACGCCCCCGG-- 458
 QY 1358 cccacctctgagatggtcccaagtcatcagcagtggtgtaccacct-----agca 1405
 Db 458 -GCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCACAGGCGCTCGGAG 516
 QY 1406 cttactcggccgatatgccaatggggggcccgagagatggaaaagaaacgatcctc 1465

Db 517 GGACACCCCGCGCTCCACCACTCGGCCAGGACAGGCGCGTGAGTACCTGCTGCTC 576
 QY 1466 ctgccagagaagatcttgccttgaacagaggtctgtcccccataaacaagtgccagagac 1525
 Db 577 TCCAAGGCCCAAGTTGGTGGCCCTCGGAGCGGAGGCGTCCCGGAGCAACAGCTGCCAAGAC 636
 QY 1526 tcacagacacccagacagcaaccacgaggatc-----gccaacatctctaccagcaaacg 1579
 Db 637 TCCACGGACACCGAGAGCAACAACAGGAGCAGCGCGTCTTATCTACCTGACCAAC 696
 QY 1580 cactgtgtctctccccagggcccgcaatggatgcctctctgaagaggtccctcgtct 1639
 Db 697 CACATGCCCGCAGCGCGCAACGCGTGTCTG-----CTCAAGGAGGAGCACCAGCGCC 747
 QY 1640 ttggaactcctcaagccccctcccatctgctgagggaactccatcaaaagtatcaacaaa 1699
 Db 748 TAGGACCTGTGCGCGCGCGCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCAC 807
 QY 1700 gaaggggaggtgatggtgttttcgatgtgaccaactgcacgtccctctctctagattat 1759
 Db 808 AGCGGGGAGCAGATGAAGGTGTACAAAGTGCGAACACTGCCGGGTGCTCTTCTTGGATCAC 867
 QY 1760 gtaatgttccacatccacatgggtgccatggttccgtgattcccttgaagtgaacatg 1819
 Db 868 GTCATGTACACCATCCACATGGGCTGCCAGCGCTCCGATGATCTTTTGTAGTCAACATG 927
 QY 1820 tgtggctatcgaagccacgacgtcgtatgattctctctcacatcgccagagagagagac 1879
 Db 928 TCGGCTTACCACAGCCAGGACCGTACGAGTTCTCTGTCGCACATAACGCGAGGAGCAC 987
 QY 1880 ag 1881
 Db 988 CG 989

Search completed: November 6, 1999, 02:38:40
 Job time: 18371 sec

Db 1447 TATCAGCCAGGACAGGTACGAGTTCTCATCCATATCAGCGGGGGAGCATCGTTAC 1506

Qy 1886 atgttgagtgagc 1899

Db 1507 CACCTGAGCTAAC 1520

RESULT 14

V65968 ID V65968 standard; cDNA; 1788 bp.

AC V65968: 14-JAN-1999 (first entry)

DE Marine Ikaros encoding cDNA mik-2.

KW C33-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;

KW differentiation marker; Immune system; corpus striatum; AIDS;

KW Alzheimer's disease; ss.

OS Mus sp.

Key Location/Qualifiers

FT Key 223..1518

CD5 /tag= a

Fi /product= "mik-2"

Fi /transl_except= (pos:385..387,aa:Gln)

PN US5824770-A.

PN 20-OCT-1998.

PN 05-JUN-1995; 465590.

PF 02-MAY-1994; US-238212.

PR 14-SEP-1992; US-946233.

PK 14-SEP-1993; US-121438.

PK 05-JUN-1995; US-465590.

PA (GEO) GEN HOSPITAL CORP.

PI Georgopoulos K;

PI WPI: 98-582821/49.

DR P-PSDB; W72671.

PT Ikaros poly:peptide(s) - useful for treating disorders of immune

PT system or corpus striatum

PS Disclosure: Column 51-56; 11pp; English.

CC The present invention describes a purified peptide having at least one

CC of the following properties: (a) it stimulates transcription of a DNA

CC sequence under the control of a delta A element, an NFkB element or an

CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of

CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide

CC consensus sequence; (c) it competitively inhibits the binding of a

CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB

CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it

CC competitively inhibits Ikaros binding to Ikaros responsive elements; or

CC (e) it inhibits protein-protein interactions of transcriptional complexes

CC formed with naturally occurring Ikaros isoforms. The proteins, provided

CC that they stimulate gene transcription under the control of delta A

CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to

CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

CC competitively inhibit binding of naturally occurring Ikaros isoforms to

CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or

CC inhibit protein-protein interactions of transcriptional complexes with

CC naturally occurring Ikaros isoforms, can be used to treat immune system

CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.

CC Alzheimer's disease. The present sequence encodes a specifically

CC claimed mouse Ikaros protein.

SQ Sequence 1788 bp; 437 A; 486 C; 495 G; 370 T;

Query Match

Best Local Similarity 17.58; Score 346.4; DB 1; Length 1788;

Matches 715; Conservative 0; Mismatches 376; Indels 63; Gaps 8;

Qy 794 ggcgaacgccgttccagtgtaacagtcgagggcatttttactcagaagtaaacctc 853

Db 382 GGTGAACGGCTTTCCAGTGCACAGTCTGGGGCCCTTCTTACCCAGAAAGGCAACCTC 441

Qy 854 ctccgtcatattaacatgcacacgggggaaacaccttttaagtctacacctcgaactac 913

Db 442 CTGCGGCACATCAGCTGCACTCGGGTGAGAGCCCTTCAATGCCATCTTGCACATAT 501

Qy 914 gcatgccaaagagagatcgctcagcgagacaccttagacacattctgtgagaaagcgg 973

Db 502 GCGTCCGCGGGAGGACGCCCTCACCAGCCACCTGAGGACGACCTCCGTTGGTAAGCCT 561

Qy 974 tacaagtgtgagttctgcggaagaagactacaagcagagaagctccctggagagacacaag 1033

Db 562 CACAAATGTGGATATTGTGCGCGGAGCTATAAACAGCGAAGCTCTTTAGAGAGGACATAAA 621

Qy 1034 gaacgctccgagagcttttctcagaacacctgacctgggggagcgtctcgaagtgtg 1087

Db 622 GAGCGATGCCACAACCTACTTGGAAAGCATGGGCCCTCCGGGCGTGTGCCCGCATTAAG 681

Qy 1088 gaggcgaagacacatcaaaag-----ccgagatgggaagtggagagagct 1129

Db 682 GAAGAACTAAACACACACAGAGATGGCAGAAAGACCTGTGCAAGATAGGACGAGAGGTC 741

Qy 1130 ctccgtctgcagagattagcaagcaatgtgtaagcgaagaaagctcgatgctcagaaa 1189

Db 742 CTGTCTCTGGACAGGCTGGCAAGCAATGTGCCAAACGTAAGAGCTCTATGCTTCAGAAA 801

Qy 1190 ttcatcggtgagaagcggcactgtcttcgagtcgaactacaatcccggtctacatgtacgag 1249

Db 802 TTCTTTGGAGACAGTGCCCTGTGACACATGCCCTATGACAGTCCCACTATGAGAAGGAG 861

Qy 1250 aaggagaacagatgatgcagaccggatgtagccaccagccatcaataacgccatcagc 1309

Db 862 -----GATATGATGACATCCACGCTGATGGACAGGCCATCAACAATGCCATCAAC 912

Qy 1310 tatctagggctgaagccttcgccccttagtccagactccgctgctccaccctctgag 1369

Db 913 TACCTGGGGCTGAGTCCCTGCGCCCATTTGGTGACAGACACCCCCCGT---AGCTCCGAG 969

Qy 1370 atggtcccatcatcagcagtggtgtaccgccatagcaacttactcgggcccagatgtccaatg 1429

Db 970 GTGGTGCCAGTCATCAGTCCATGTACCAGCTGCACAAAGCCCCCTCAGATGSCCCCCCA 1029

Qy 1430 ggg-----gcccgcagagagatggaaaagaacacgcatcctcctgccagagag 1477

Db 1030 CGGTCCCAACCATTCACACAGGACGCGGTGGATAACTTGTGCTGTCTCCAAAGGCCAAG 1089

Qy 1478 atcttgctcttgaaagaggtctgtcccccataacagtgcccagagactccacacaccc 1537

Db 1090 TCTGTCTCATGGAGGAGAGGCGCTCCCGAGCAACAGCTGCCAAGACTCCACAGATACA 1149

Qy 1538 gacagcaacacagag---atcgccaacatctctaccagcaaacgacagcagtggtcctcccc 1594

Db 1150 GAGAGCAACGCGAGGACAGCGCGCGCTTATCTACTTAACCAACCAACATCAACCCG 1209

Qy 1595 caggccgcaatggagtgctcttctgaaggaggtccctcgtctctttgaactcctcaag 1654

Db 1210 CATGCACGCAATGGGCTGGCTCTC---AAGGAGGAGCAGCGGCGCTTACGAGGTGCTGAGG 1266

Qy 1655 cccctctcccatctgctcagagactccatcaagtgatcaacaagaaggaggtgatg 1714

Db 1267 GCGGCTCTAGAAACCTCGCAGGATGCCCTTCCGTGTGGTGAGCAGAGTGCAGCAGCTG 1326

Qy 1715 gatgtgtttcctgatgaccactgcaacgtccctctctctctctctctctctgtgtgttcacatc 1774

Db 1327 AAGGTGTACAAGTGCAGAACACTGCGCGGTCTCTCTCTGGATCACGTCATGTATACCAT 1386

Qy 1775 cacatg-----gggtgccatggttcccgatgaccccttttgatgtgaactgtgtg 1825

Db 1387 CACATGGGTGCCATGGCTGCCATGGCTTTTCGGGATCCCTTTTGTAGTGTAAACATGTGTGT 1446

Qy 1826 tatcgagccacagatcgctatgattctcctctcctcctcctcctcctcctcctcctcctcctc 1885

Db 1447 TATCAGCCAGGACAGTACGAGTGTCTCTCCATATCAGCGGGGGAGGATCGTTAC 1506

Qy 1886 atgttgagtgagc 1899

Db 1507 CACCTGAGCTAAC 1520


```

Db 969 AGAAGACCTGTCCAGATAGGACGAGAGGTCCTTGTCTGGACAGCTGGCAAGCAA 1028
Qy 1156 tggctgaagcaaaagctcgcctcagaaattcatcgtgagaagcgcaactgttt 1215
Db 1029 TGTGCCAAACGTAAGAGCTCTATGCTCAGAAATTTCTTGAGACAAGTGCCTGCAGA 1088
Qy 1216 cgtatccaaactacatcccggtacatgtacgagaagagacagatgatgcagaccg 1275
Db 1089 CATGCCCTATGACGTGCAACTATG-----AGAAGGAGGATATGATGACATCCCA 1139
Qy 1276 gatgatgacaaacgcatcaataacgcccagctataggggtgaagccttcgcgcc 1335
Db 1140 CGTGATGACCAAGGCCATCAACAATGCCATCAACTACCTGGGGGTGAGTCCCTCGGCC 1199
Qy 1336 cttagtcagactccgctcctccacactctgagatggtcccaagtcacagcagtgta 1395
Db 1200 ATTGGTGACAGACCCCGGT--AGTCCGAGGTGGTGGCAGTCATCAGTCCATGTA 1256
Qy 1396 cccatagcacttactcgggcccagatatgccaatggggccc-----ccgcagga 1443
Db 1257 CCAGCTGCAAGACCCCTCTAGATGGCCCCCAGCGTCCAAACCATTCAGCACAGGAGCG 1316
Qy 1444 gatgaaagaaacggtcctcctccagagagagatcttgcttgaagcaggtcttc 1503
Db 1317 CGTGATTAATCTGTGCTGTCTCAAGGCCAAGTCTGTGTCTATCGGAGCGAGGCGCTC 1376
Qy 1504 ccccaataacagtcgagagctccacagacacgacagcaaccagagg---atcgcca 1560
Db 1377 CCCGAGCAACAGCTGCCAAGACTCCACAGATACAGAGCAACGCGGAGGACAGCGAG 1436
Qy 1561 acatctaccgcaaaagccagtcgttctcccccagggccgcaatggagtcctcttt 1620
Db 1437 CGGCTTATCTACCTAAGCAACACATCAACCCGATGATGCGCAATGGCTGGC---TCT 1493
Qy 1621 gaagaggtcctcgtcttttgaactcctcaagccccctccatctcgtcagggactc 1680
Db 1494 CAAGGAGGAGCAGCGGCTACGAGGTGCTGAGGGCGGCTCAGAGAACTCGCAGAGTGC 1553
Qy 1681 catcaagtgtacaaagaggggagtgatgattgttctgattgtgacctgcca 1740
Db 1554 CTTCCGTGTGTGTCAGCAGAGTGGCGAGCAGCTGAAGTGTACAGTGCGRACACTGCGG 1613
Qy 1741 gctcctctcctagattatgtgatttccacatccacatg-----gggtgcatgg 1791
Db 1614 CGTGCTCTCTGTGATCAGCTCATGTATACCATTCACATGGGCTGCCATGGCTGCATGG 1673
Qy 1792 ttccgtgatcccttgatgttaacatgtgtgctatgaagcagacatcgctatgatt 1851
Db 1674 CTTTCGGATCCCTTTGAGTGTAAATGTGTGTGTATCAGACCCAGGACAGGTACGAGTT 1733
Qy 1852 ctctctcacatcgccagagagagacagagccattgttgaagtgaac 1899
Db 1734 CTCATCCCATATCACGCGGGGGAGCATCGTTTACCACCTGAGCTAAAC 1781

RESULT 11
V42805
AC V42805 standard; cDNA; 1788 BP.
DE 11-JAN-1999 (first entry)
KW Mouse Ikarois isoform mik-2 cDNA.
KW Ikarois; mik-2; transcription factor; mouse; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
KW Alzheimer's disease; therapy; diagnosis; ss.
OS Mus sp.
FH Key
FT Location/Qualifiers
FT 223..3518
FT /*tag= a
FT /transl_except= (pos:385,,387, aa:Gln)
FN CA2194256-A.
PD 05-MAR-1998.
PF 02-JAN-1997; 194256.

```

```

PR 05-SEP-1996; US-711417.
PA (GEO ) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR P-PSDB; W70963.
PT New nucleic acid encoding Ikarois protein involved in early
PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer and
PT to control cell differentiation
PS Claim 7; Page 68-70; 158pp; English.
CC This is the nucleotide of mouse Ikarois cDNA (isoform mik-2) that
CC codes for a 431-amino acid zinc finger protein (see W70963) that is
CC involved in the early differentiation of lymphocytes. A cDNA
CC library constructed from the mature murine T cell line E14 was
CC screened with a multimerised oligonucleotide (see V42829) derived
CC from a protein binding site (see V42804) of the CD3-delta enhancer
CC to identify T cell specific sequences that bind and mediate
CC enhancer function. An isolated clone was designated Ikarois and
CC contained the 1788 bp sequence. A 300 bp 3'-terminal segment of
CC this was used to identify 4 other Ikarois isoforms (see V42807-10).
CC Different isoforms arise by differential splicing of Ikarois gene
CC transcripts. Isoform mik-2 lacks exon 3. It is expressed in
CC embryonic liver, thymus and brain, but only in thymus and spleen
CC after birth. The Ikarois gene is located at the proximal arm of
CC murine chromosome 11. Ikarois proteins are suggested to play a role
CC as a genetic switch regulating entry into the T cell lineage. The
CC murine and human (see V42806, V42811 and V42840) Ikarois sequences
CC are very similar. The invention provides Ikarois nucleic acids,
CC vectors and host cells expressing Ikarois proteins. These are used
CC to treat T and B cell diseases (e.g. immune deficiencies caused by
CC drugs, radiation or cancers), to control expression of heterologous
CC genes placed under control of an Ikarois-responsive element, to
CC treat nervous system diseases (e.g. Alzheimer's disease) and to
CC modulate cell division, amplification or differentiation, especially
CC in haematopoietic cells. Some Ikarois isoforms are antagonistic of
CC others and may be used to inhibit interaction with DNA sequences.
CC The same effect can be achieved with Ikarois-binding oligonucleotides.
CC Examining the expression of the Ikarois gene, or its allelic
CC structure, can be used to assess risk of acquiring the above
CC diseases.
SQ Sequence 1788 BP; 437 A; 488 C; 493 G; 370 T;

Query Match 17.5%; Score 348; DB 1; Length 1788;
Best Local Similarity 62.0%; Pred. No. 5.5e-78;
Matches 716; Conservative 0; Mismatches 375; Indels 63; Gaps 8;

Qy 794 ggcgaagcccccagtgtaatacagtcgaggggcacatcttttactcagaaggaacctc 853
Db 382 GGTGAACGGGCTTTCCAGTGCACACAGCTCTGGGGCTCTTTTACCAGAAAGGCAACCTC 441
Qy 854 ctccgtcatataaactgcacacgggggaaaaacaccttttaagtctcacctctgcaactac 913
Db 442 CTGGCGGACATCAAGCTGCACCTCGGTGAGAGCCCTTCAATGCCATCTTTGCACTAT 501
Qy 914 gcatgccaagagagatgcgctcacgggacacaccttaggacacattctgtggagaagcgg 973
Db 502 GCCTGCCCGCGGAGGAGCGCCCTCACCGGCACCTGAGGAGCGCACTCCGTGGTAAAGCCT 561
Qy 974 tacaagtgtgagttctcggagaagaagctacaagcagagaagctccctcgagagacacaag 1033
Db 562 CACAAATGTGGATATTGTGGCGGAGCTATAAACAGGCAAGCTCTTTAGAGAGGACATAA 621
Qy 1034 gaacgctgcccagcttttcttcagaacccctgacctgggggacg-----ctgcaagtg 1087
Db 622 GAGGATGCCACAACACTACTTGGAAAGCATGGGCTTCGGGGCGTGTGCCCATGCTATTAG 681
Qy 1088 gaggcaagacacatcaaaag-----ccgagatgggaagtgaagagact 1129
Db 682 GAAGAAATCAACCAACAGAGATGGCAGAGACCTGTGCAAGATAGGAGCAGAGAGGTCC 741
Qy 1130 ctgctctgacagattagcaagcaatgtggttaagcgaaaaaagctcgatgcctcagaaa 1199

```

```
Db 1200 ATTGGTCAGACACCCCGCGT---AGCTCCGAGGTTGGTCCAGTCATCAGCTCCATGTA 1256
Qy 1396 cccatagcacttactcggccgatgccaatggggcc-----ccgcagga 1443
Db 1257 CCAGCTGCACAAGCCCTCAGATGGCCCCCACCAGTCCACACCTTACGACAGGAGCC 1316
Qy 1444 gatgaaagaaacggatcctcctccgagagaagatcttgccttctgaacgggtctgtc 1503
Db 1317 CGTGGATAACTTGTCTGCTCTCTCCAAAGCCCAAGTGTGTATCGAGGAGAGGCGCTC 1376
Qy 1504 cccataaacagtgcagagactccacacacacacacacacacacacacacacacacacac 1560
Db 1377 CCGGACACAGCTGCCAGATCTCCACATACAGAGACACGCGGAGAACAGCGCAG 1436
Qy 1561 acatctctaccagcaaacgacgtggtctcccccagggcccaatggatgctctcttct 1620
Db 1437 CGGCTTATCTACCTAACCAACACATCAACCCGCATGCACGCAATGGCTGGC---TCT 1493
Qy 1621 gaaggaggtccctcgtcttttgaactctcctcaagccctcccatctgctgaggaactc 1680
Db 1494 CAAGGAGGAGCAGCGCTACGAGTGTCTGAGCGCGCTCAGAGAACTCGCAGATGC 1553
Qy 1681 catcaaatgatcaacaagaagggagtgatggtgtgttctgagtgacactgcca 1740
Db 1554 CTTCCTGTGGTGCAGCACAGTGGCAGCAGCTGAAGTGTACAGTGCACACATGCCG 1613
Qy 1741 cgtcctctctctagattatgtgatgttaccatccacatg-----gggtgcctgg 1791
Db 1614 CGTCTCTTCTGGATCAGCTCATGTATACCATTCATGCGGTGCGTGCATGG 1673
Qy 1792 tttcctggtatccctttgaatgaatacattgtgtggtctatcaagccacgactgagtt 1851
Db 1674 CTTTCGGGATCCCTTTGAGTGTAAACATGTGTGTATATCAGCCAGGACAGGTACGAGT 1733
Qy 1852 ctctctccatcgcagagagagacacagaccatgttgtaagtgaac 1899
Db 1734 CTCATCCCATATCACCGCGGGGAGCATCTTACCACCTGAGCTAAAC 1781
```

RESULT 10

```
V42808
ID V42808 standard; cDNA; 2049 BP.
AC V42808:
DT 11-JAN-1999 (first entry)
DE Mouse Ikaros isoform mik-1 cDNA.
KW Ikaros; mik-1; transcription factor; mouse; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
OS Alzheimer's disease; therapy; diagnosis; ss.
FH Mus sp.
FT Key Location/Qualifiers
FT CDS 223..1779
FT exon /*tag= a
FT exon 223..384
FT exon /*tag= b
FT exon /number= Ex1/2
FT exon 385..643
FT FT /*tag= c
FT FT /number= Ex3
FT FT 644..810
FT FT /*tag= d
FT FT /number= Ex4
FT FT 811..933
FT FT /*tag= e
FT FT /number= Ex5
FT FT 934..1076
FT FT /*tag= f
FT FT /number= Ex6
FT FT 1077..1779
FT FT /*tag= g
FT FT /number= Ex7
PN CA2194256-A.
PD 05-MAR-1998.
```

Query Match 19.1%; Score 379.6; DB 1; Length 2049;
Best Local Similarity 61.8%; Pred. No. 7e-86;
Matches 771; Conservative 0; Mismatches 414; Indels 63; Gaps 8;

Qy 700 cagcagaccacacagtggaagatgaactgcagctgctgcgggttctctgctgattgctt 759
Db 549 CATTCGACTTCTTAACGGAAACAAAGTGTATATCTGTGGATCGTTTGCATCGGCC 608
Qy 760 caacgtctgtgatttcataagcgaagccataccgcgcgaacgcgcgttccagtgtaata 819
Db 609 CAATGTGCTCATGTGTTCAAAAGAAAGTCACTACTGTGTAACGGCTTTTCCAGTGAACA 668
Qy 820 gtgcggggcattctttactcagaaggttaacctctcctcgtcatattaaactgcacacgg 879
Db 669 GTCTGGGGCTTCTTTACCCAGAAAGGCAACCTCTCGCGGCACATCAAGCTGCATCTCGG 728
Qy 880 ggaataaaccttttaagtgtcacctctgcactgcagctgcacacagagagagatgagctcac 939
Db 729 TGAGAAAGCCCTTCAATGCCATCTTTGCACTATGCTCTGCCCGGAGGAGCGCCTCAC 788
Qy 940 gggacacattagacacattctgtggagaagccgttacaagtgtgagttcttcggaagaag 999
Db 789 CGGCCACCTGAGGAGCGCACTCCGTGTGTAAGCCTCACAAATGTGATATATTGCGCGGAG 848
Qy 1000 ctacaagcagagaactccctgggagagacacaaaggaacgcgtgcgcgagcttttcttcagaa 1059
Db 849 CTATTAACAGCGAAGCTTTTAGAGAGACATTAAGAGCGATGCCCACTACTTCTGGAAG 908
Qy 1060 cctgtacctgggggagc-----ctgcaagtgtggaggaagacacacatcaaac----- 1107
Db 909 CATGGCCCTTCGGGCGTGTGCCCAGTCTATTAAAGGAAGAACTAACCAACAGGATGGC 968
Qy 1107 -----ccgagatgggaagtgaagagctctcctgacagattagcaagca 1155

```

Db 1140 CGTATGACGACGACATACATATGCAATCACTACCTGGGGCTGAGTCCCTGCGGCC 1199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1336 cttagtcagactccgctcgtccacactcgtgagatggtccagtcacagcagtgta 1395
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1200 ATTGGTCAGACACCCCGGT---AGTCCGAGGTGTGTCAGTCATCAGCTCCATGTA 1256
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1396 cccatagcacttactcgggcccgtatgcaaatgggggccc-----ccgagga 1443
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1257 CCACTGCACAAAGCCCGCTCAGATGGCCCGCCACGCTCCAAACATTCAGCACAGGACG 1316
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1444 gatgaaagaaacggtccctcctgcccagagaaatcttcttcttaaacagtgctgc 1503
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1317 CTGGATAAATTGCTGCTGTCTCAAGGCCAAAGTGTGTGTCATCGGAGCGAGGCGCTC 1376
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1504 ccccaataacagtgcccaggactccagacacacagcagcagcaaccacaggg---atcgcca 1560
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1377 CCCGAGCAACAGCTGCCAAGACTCCACAGATACAGAGAGCAACGCGGAGGAACAGCGAG 1436
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1561 acatctacagcaaacgacgtggtctctcccccagggcccgccgaatgggagtccttct 1620
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1437 CGGCTTATCTACCTAAACCAACCATCAACCCGCTATGCAGCAATGGGCTGGC---TCT 1493
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1621 gaaggagtcctcgtctcttttgaaactcctcaagcccccctccatctgcctgagggactc 1680
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1494 CAAGGAGGAGAGCGCGCTACGAGGTGCTGAGGGCGGCTCAGAGAACTCGCAGGATGC 1553
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1681 catcaagtgatcaacaagaaggggagtgatggtgtgttctgagtgaccactgcca 1740
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1554 CTTCGCTGTGTACGACGAGTGGCGAGCAGCTGAAGGTGTACAAAGTCCGAACACTGCCG 1613
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1741 gctcctctcttagattatgtgattcaaccatccaatg-----gggtgccatgg 1791
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1614 CGTCTCTTCCTGTGATCACGTATGATATACATTACATTCATGCGCTGCCATGG 1673
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1792 ttccgtgatccctttgagtgataacatgtggtctatcgaaagccacacatgcctatgatt 1851
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1674 CTTTCGGGATCCCTTTGAGTGTACATGTGTGTTATACAGCCAGGACAGGTACGAGTT 1733
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1852 ctctctcacatccagagagagacagagccatgttgagtgagc 1899
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1734 CTCATCCATATACGCGGGGGAGCATCGTTACCACCTGAGCTTAAC 1781
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 9

```

V6971
ID V6971 standard; cDNA; 2049 BP.
AC V6971;
DE 14-JAN-1999 (first entry)
DE Mouse Ikaros encoding cDNA mik-1.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; immune system; corpus striatum; AIDS;
KW Alzheimer's disease; ss.
OS Mus sp.
FH Key Location/Qualifiers
FT CDS 223..1775
FT FT /*tag= a
FT FT /product= "mik-1"
FN US5824770-A.
PD 20-OCT-1998.
PF 05-JUN-1995; 465590.
PR 02-MAY-1995; US-238212.
PR 14-SEP-1992; US-946233.
PR 14-SEP-1993; US-121438.
PR 05-JUN-1995; US-465590.
PA (GPHO ) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-582621/49.
DR P-PSDB; W72674.
DR Ikaros poly:peptide(s) - useful for treating disorders of immune
PT system or corpus striatum
PS Disclosure: Column 61-66; 111pp; English.
CC The present invention describes a purified peptide having at least one

```

```

CC of the following properties: (a) it stimulates transcription of a DNA
CC sequence under the control of a delta A element, an NFkB element or an
CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
CC consensus sequence; (c) it competitively inhibits the binding of a
CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC (e) it inhibits protein-protein interactions of transcriptional complexes
CC formed with naturally occurring Ikaros isoforms. The proteins, provided
CC that they stimulate gene transcription under the control of delta A
CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit binding of naturally occurring Ikaros isoforms to
CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC inhibit protein-protein interactions of transcriptional complexes with
CC naturally occurring Ikaros isoforms, can be used to treat immune system
CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC Alzheimer's disease. The present sequence encodes a specifically
CC claimed mouse Ikaros protein.
CC Sequence 2049 BP: 514 A; 531 C; 574 G; 430 T;
SQ

```

```

Query Match 19.1%; Score 379.6; DB 1; Length 2049;
Best Local Similarity 61.8%; Pred. NO. 7e-86;
Matches 771; Conservative 0; Mismatches 414; Indels 63; Gaps 8;

```

```

Qy 700 cagcagacacacacagtggaagatgaactcgacgtgctgctggttatcctgcagctt 759
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 CATTGACTTCTTACGGAAACTAAAGTGTGATATCTGTGGGATCGTTTGCATCGGGCC 608
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 760 caacgtcttgatggttcataagcgaagccataccgcgaacgcccttccagtgtaatca 819
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 609 CAATGTCTCATGGTTCACAAAAGAGTACTACTGTGAACGGCTTTTCAGTGCAACCA 668
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 820 gtgcggggcactctttactcagaaaggttaacctcctcctcgtcattataaactgcacacgg 879
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 669 GTCTGGGGCTCTCTTACCAGAAAGGCAACCTCCTCGGCACATCAAGCTGACTCGGG 728
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 880 ggaataacacctttaaagtgtcacctctgcaactacgacatgccaaagagagatcgctcac 939
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 729 TGAGAACCCCTTCANATGCCATCTTTGCAACTATGCTGCCCGCGGAGGAGGCCCTTAC 788
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 940 gggacacaccttagacacattctgtggaagccgtacaaagtgtgagttctgcggaagaag 999
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 789 CGGCCACCTGAGGACGCACTCCGTTGGTAAGCCTCACAAATGTGGATATTGTGCCCGAG 848
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1000 ctacaagcagagaagctccctggagagacacaaggaacgctgcgcgagcttttttcagaa 1059
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 849 CTATAACACGCAAGCTCTTTAGAGAGGACATAAAGAGCGATGCCACACTACTTTGGAAG 908
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1060 cctgacactggggagc-----ctcaagtgtggaggaagacacacatcaag----- 1107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 909 CATGGGGCTTCCCGGGCGTGTGCCACAGTCATTAAAGGAAGAACTAACCAACACAGATGCG 968
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1107 -----ccgagatgggaagtggagagctctcgtctgagacagattagcaagcaa 1155
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 969 AGAAGACCTGTGCAAGATAGGACAGAGAGTCCCTTGTCTTGACACAGGCTGCAAGCAA 1028
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1156 tgtggtaagcgaagaagctcgtgctcagaaattcattcgttgagaagcggcactgctt 1215
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1029 TGTCGCCAAACGTAAGAGCTCTTATGCTCTCAGAAATTTCTTTGGAGACAAGTGCCTGTGAGA 1088
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1216 cgatgccaactacaatcccggtctacatgtacgagagagagaacagatgatgcagaccgg 1275
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1089 CATGCCCTATGACAGTGCCCAACTATG-----AGAGGAGGATATGATGACATCCCA 1139
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1276 gatgatggacaaagccatcaataacgccatcagctacttagggggctgaaagccttcgccc 1335
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1140 CGTGATGACGACGCGCATCAACAATGCCATCAACTACCTGGGGCTGAGTCCCTGGGCC 1199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1336 cttagtcagactccgctcgtctccacctctgagatggttccccagtcacagtgtagta 1395
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db 933 CGCGTGGAGTACCTGCTGCTCTCTCCAGGCCAAGTTGGTGCCCTCGGAGCGGAGGC 992
Qy 1501 gtcccccaataacagtccaggactccacagacacccagacacccagcaggatc----- 1557
Db 993 GTCCCCGGAGCAACAGCTGCCAAGACTCCACGAGACACCGAGACAAACAGGAGGACGCG 1052
Qy 1557 --gcaacatctctaccagcaaacagcagtggtctctccccagccccaatggatgcc 1614
Db 1053 CAGCGGTCTTATCTACCTGACCAACACATCGCCGCGAGCGCGAACCGGTGTCG----- 1108
Qy 1615 tcttctgaaggaggtccctcgctcttttgaactctctcaagcccccctccatctgctgag 1674
Db 1108 ---CTCAGGAGGAGCACCAGCGCTCAGCACTGTGCGCGCGCTCCGGAACTCGCA 1163
Qy 1675 ggaatccatcaagtgtatcaacaaagagggaggtgtagtgtgtttcgtatgtgacca 1734
Db 1164 GGACGCGTCCGCGTGGTCAGCACCGCGGAGCAGATGAAGGTGTACAAGTGGGAACA 1223
Qy 1735 ctgcaacgtctctcttagattatgtatgtttcaaccatcacatggggtgcccattggtt 1794
Db 1224 CTGCGGGGTGCTTCTCCCTGGATCAGCTCATGTATACACCATCCACATGGCGTCCACGGCTT 1283
Qy 1795 ccgtgatccctttagtgaacatgtgtggtatctgagcagccacgacgtcgtatgattctc 1854
Db 1284 CCGTGATCTTTGAGTGCACATGTGCGGTACCCACAGCCAGGACCGGTACGAGTTCTC 1343

Qy 1855 ctctcacatgccagagagagacag 1881
Db 1344 GTCCGACATAACCGGAGGGAGCACCG 1370

RESULT 7
ID V42840
AC V42840;
DT 11-JAN-1999 (first entry)
US Human Ikaros isoform hlk-1 cDNA.
KW Ikaros; hlk-1; transcription factor; human; lymphocyte;
  cell differentiation; T cell; cancer; immunodeficiency;
  Alzheimer's disease; therapy; diagnosis; ss.
NC Homo sapiens.
FH Key
FT exon 1..26
FT /tag= a
FT /number= Ex1
FT 27..163
FT /tag= b
FT /number= Ex2
FT 164..420
FT /tag= c
FT /number= Ex3
FT 421..588
FT /tag= d
FT /number= Ex4
FT 589..714
FT /tag= e
FT /number= Ex5
FT 715..849
FT /tag= d
FT /number= Ex6
FT 850..1551
FT /tag= e
FT /number= Ex7
FT CA2194256-A.
PN 05-MAR-1998.
PD 02-JAN-1997; 194256.
PR 05-SEP-1996; US-711417.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR P-PSDB; W70971.
PT New nucleic acid encoding Ikaros protein involved in early
```

```

PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer, and
PT to control cell differentiation
PS Claim 1; Page 127-129; 158pp; English.
CC This is the nucleotide of human Ikaros cDNA (isoform hlk-1) that
CC codes for a 516-amino acid zinc finger protein (see W70971) that is
CC involved in the early differentiation of lymphocytes. It was
CC isolated from a Jurkat T cell line library using mouse Ikaros exon
CC 7 cDNA as probe. The Ikaros gene maps between p11.2-p13 on human
CC chromosome 7. The human and murine Ikaros sequences (see V42805-11
CC and V42840) are highly conserved. Differential splicing of Ikaros
CC gene transcripts gives rise to different Ikaros protein isoforms.
CC The invention provides Ikaros nucleic acids, vectors and host cells
CC expressing Ikaros proteins. These are used to treat T and B cell
CC diseases (e.g. immune deficiencies caused by drugs, radiation or
CC cancers), to control expression of heterologous genes placed under
CC control of an Ikaros-responsive element, to treat nervous system
CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
CC amplification or differentiation, especially in haematopoietic
CC cells. Some Ikaros isoforms are antagonistic of others and may be
CC used to inhibit interaction with DNA sequences. The same effect
CC can be achieved with Ikaros-binding oligonucleotides. Examining
CC the expression of the Ikaros gene, or its allelic structure, can be
CC used to assess risk of acquiring the above diseases. 265 T;
SQ Sequence 1551 BP; 393 A; 450 C; 443 G; 265 T;

Query Match 20.8%; Score 411.8; DB 1; Length 1551;
Best Local Similarity 63.3%; Pred. No. 5.5e-94;
Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

Qy 700 cagcagaccacagtggaagatgaactgcgacgtgtgcgggttatccgtgactgctt 759
Db 327 CATTCGACTTCTTAACGGAAACTAAAGTGTGTATCTGGGATCATTTGCATCGGGCC 386
Qy 760 caactcttgatgttcataaggaagccatacgcgggaagccgcgttcagtgataatca 819
Db 387 CAATGTGCTCATGTTTCCAAAAGAACACACACCTGGAGAACGGCCCTTCCAGTGCATCA 446
Qy 820 gtgcgggacatcttttactcagaagagtaacctctcctcgtcatattaaactgcacaggg 879
Db 447 GTGCGGGCCCTCATTTACCCAGAGGGCAACCTGCTCCGGCACATCAAGTGTCTCCGG 506
Qy 880 ggaacaccccttttaagtgtcacctctgcacgtacgagtgccaaaggagagatgcgtcac 939
Db 507 GGAGAAGCCCTTCAATGCCACCTCTGCAACTACGCCCTGCCCGGAGGAGCGCCCTCAC 566
Qy 940 gggacaccttaggcacacattctgtggagaagccgtacaaagtgtgagttctgcggaagaag 999
Db 567 TGGCCACCTGAGGACGACACTCCGTTGTTAACTCAAAATGTGGATATTGTGCCCGGAAG 626
Qy 1000 ctacaagcagagaagctcccttgaggagcacaaggaacgctgcccagcttttcttcagaa 1059
Db 627 CTATAAACAGCGAAGCTCTTTAGAGGAACATAAAGACGCGTGCACACACTACTTGGAAAG 686
Qy 1060 ccttgacct--gggggacgctgcaagtgtgaggagcaagacacatcaaaagc----- 1108
Db 687 CATGGGCTTCCGGGCACACTGTACCCAGTCATTTAAAGAGAAACTAAGCACAGTGAAT 746
Qy 1108 -----cgagatgggaagtgcagagctctcgtctctgcagacattagcaag 1152
Db 747 GGCAGAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGTGGACAGACTAGCAAG 806
Qy 1153 caatgttgctaagcgaaaaagctcgatgcctcagaaattcatcgttgtagaagcgccactg 1212
Db 807 TAATGTGCCAAACGTAAGAGCTCTATGCTCAGAAATTTCTTGGGCACAAGGCCCTGTC 866
Qy 1213 cttcgatgccaaactacaatccccggctacatgtacgagaagagagacgatgatgcagac 1272
Db 867 CGACACGCCCTACGACAGTGC-----CAGGTACGAGAAGGAGAACGAAATGATGAAGTC 920
Qy 1273 ccgcatgatggaccaagccatcaataacgcatctatctaggggtgaagccttccg 1332
```

Db 993 GTCCCCGAGCAACAGCTGCCAACACTCCACGGACCGGACCGAGAGCAACACGAGGAGCAGCG 1052
 Qy 1557 --gcaacatctctaccagcaagccacgtggtctctccccagccgcgaatgggagtc 1614
 Db 1053 CAGCGGTCTTATCTAGCTACCAACACATCGCCGCGGCAACGCGGTGTGCG----- 1108
 Qy 1615 tctctgaaggaggtccctcgctcttttgaactcctcaagccctccctcctctgctgag 1674
 Db 1108 ----CTCAGGAGGAGACCGCGCTTACGACCTGCTGCGCGCGCTCCGAGAACTCGGA 1163
 Qy 1675 ggactccataaagtatcaacaaagggaggtgagtggtgttttcgagtgaacca 1734
 Db 1164 GGACGGCTCCGCGTGTGTCAGCACCAGCGGGGAGCAGATGAAGTGTACAAGTGGCA 1223
 Qy 1735 ctgcaagctctctctagattatgtgttcaaccatccacatgggtgcccattgttt 1794
 Db 1224 CTGCCGGGTGCTCTCTGGATCAGCTCATGTACACATCCATGGGTGCGACGGCTT 1283
 Qy 1795 ccgtgatcccttggagtgtaacatgtgtggtatcgaaagccagatcgctatgattctc 1854
 Db 1284 CCGTGATCTTTGAGTGCNACATGTGGCTTACCACAGCCAGGACCGGTACGAGTTCTC 1343
 Qy 1855 ctctcatcgccagaggagagcacag 1881
 Db 1344 GTGGCATAAACGCGAGGGGAGCACCG 1370

RESULT 6

ID V42806
 AC V42806; standard; cDNA; 1386 BP.
 DT 11-JAN-1999 (first entry)
 DE Human Ikaros isoform hik-1 cDNA.
 KW Ikaros; hik-1; transcription factor; human; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 1..255
 FT /tag= a
 FT /number= Ex3
 FT 256..423
 FT /tag= b
 FT /number= Ex4
 FT 424..549
 FT /tag= c
 FT /number= Ex5
 FT 550..684
 FT /tag= d
 FT /number= Ex6
 FT 685..1386
 FT /tag= e
 FT /number= Ex7
 PN CA2194256-A.
 PD 05-MAR-1998.
 PE 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR P-PSDB; W70964.
 PT New nucleic acid encoding Ikaros protein involved in early
 PT differentiation of lymphocytes - existing in several isoforms, and
 PT related products, used to treat e.g. immune diseases or cancer and
 PT to control cell differentiation
 PS Claim 7; Page 70-72; 158pp; English.
 CC This is the nucleotide of human Ikaros cDNA (isoform hik-1) that
 CC codes for a 461-amino acid zinc finger protein (see W70964) that is
 CC involved in the early differentiation of lymphocytes. It was
 CC isolated from a Jurkat T cell line library using mouse Ikaros exon
 CC 7 cDNA as probe. The Ikaros gene maps between pl1.2-pl3 on human
 CC chromosome 7. The human and murine Ikaros sequences (see V42805-11
 CC and V42840) are highly conserved. Differential splicing of Ikaros

CC gene transcripts gives rise to different Ikaros protein isoforms.
 CC The invention provides Ikaros nucleic acids, vectors and host cells
 CC expressing Ikaros proteins. These are used to treat T and B cell
 CC diseases (e.g. immune deficiencies caused by drugs, radiation or
 CC cancers), to control expression of heterologous genes placed under
 CC control of an Ikaros-responsive element, to treat nervous system
 CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
 CC amplification or differentiation, especially in haematopoietic
 CC cells. Some Ikaros isoforms are antagonistic of others and may be
 CC used to inhibit interaction with DNA sequences. The same effect
 CC can be achieved with Ikaros-binding oligonucleotides. Examining
 CC the expression of the Ikaros gene, or its allelic structure, can be
 CC used to assess risk of acquiring the above diseases. 234 T;
 SQ Sequence 1386 BP; 352 A; 403 C; 397 G;

Query Match 20.8%; Score 411.8; DB 1; Length 1386;
 Best Local Similarity 63.3%; Pred. No. 5.3e-94;
 Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

Qy 700 cagcagacaaaccagtggaagatgaactgcgagctgtgcgggttatcctcgttagctt 759
 Db 162 CATTCGACTTCCTAACGGAAACTAAAGTGTATCTGTGGGATCATTTGTCATCGGGCC 221
 Qy 760 caacgtttgatgttcaagaagccataccggaagccggtccgttccagtgtaacca 819
 Db 222 CAATGTCTCATGGTTCAAAAAGAACCCACACTGGAGAACGGCCCTTCCAGTGTCAATCA 281
 Qy 820 gtgcggggcatcttttactcagaaggttaacctctcctcgtcatattaaactgcacacgg 879
 Db 282 GTGCGGGGCCCTCATTCACCCAGAGGGCAACTGTCTCGGCACATCAAGCTGTGATCCGG 341
 Qy 880 ggaataacctttaaagtgtcaacctgtcaactacatcgatccaaaggagagatgcctcac 939
 Db 342 GGAGAAAGCCCTTCAAAATGCCACTCTGCAACTACGCCCTGCCGCGGAGGACGCCCTCAC 401
 Qy 940 gggacaccttagcacacattctgtgagaagccgtacaaagtgtgaagttctcggaagaag 999
 Db 402 TGGCCACCTGAGGAGCGCACTCGGTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAG 461
 Qy 1000 ctacaagcagaagaagctccctggaggagacacaaaggaacgcgtgcgagcttttcttcagaa 1059
 Db 462 CTATAACAGCGCAACGCTCTTTAGAGGAACATAAAGAGCGCTGCCACACTACTTGGAAAG 521
 Qy 1060 cccctgacct--gggggagcgtgcaagtgtggaggcaagacacatcaaaagc----- 1108
 Db 522 CATGGCCCTTCGGGCACTGTACCCAGTCAAAAAAGAAAGAACTAAGACACAGTGAAT 581
 Qy 1108 -----cgagatgggaagtgaagagctctcctcgtggacagattagaag 1152
 Db 582 GGCAGAACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGTGGACAGACTAGCAAG 641
 Qy 1153 caatgtgctaagcgaataagcgtcgtcctcagaaattcatcggtgagaagcgcactg 1212
 Db 642 TAATGTGCCCAACGTAACAGCTCTATGCCCTCAGAAATTTCTTGGGGACAAGGGCCCTGGC 701
 Qy 1213 cttcgatgccaaactacaatcccggtctacatgtacgagaaggaagacagatgacagac 1272
 Db 702 CGACACGGCCTACGACAGTGC-----CACGTACGAGAAGGAGAAACGAATGATGAAGTC 755
 Qy 1273 ccggatgatgacaaagccatcaataaacccatcagctatctaggggctggaagccttcg 1332
 Db 756 CCACGTGTATGGACCAAGGCATCAACAACGCCATCAACTACCTGGGGCGCGAGGCCCTCGC 815
 Qy 1333 ccccttagtccagactccgctcgtcccacctctgagatggtcccagtcacagcagtg 1392
 Db 816 CCGCTGTGTGACAGCGCCCCCG---GGGGTTCAGAGTGTGTCGCGGTATCAGCCCCAT 872
 Qy 1393 gtac-----cccatagcacttactcgggcccagatgccaatggggggccccgca 1440
 Db 873 GTACCAGCTGCACAGGCGCTCGAGGGGACCCCGCGCTCCAAACCACTCGGGCCAGGACAG 932
 Qy 1441 ggagatggaagaagaacggatcctcctccagagagaagattcttgccttctgaacgaggtct 1500

Db 933 CGCCGTGGAGTACCTGCTGCTCTCAAGGCCAAGTTGTTGCTCGGAGCGGAGGC 992
 QY 1501 gtcccccataacagtgcccaggagctccacagacacagacacagacacagaggtatc----- 1557
 Db 993 GTCCCCCGAGCAACAGCTGCCAAGACTCCACGACACACCGAGAGCAACAAAGAGGAGCAGCG 1052
 QY 1557 --gcaacaatctctaccagcaaacgacgctggtctctccccagggcccgcaatggatgcc 1614
 Db 1053 CAGCGGTCTTATCTACCTAGCAACCACTATCGCCGAGCGCGGCAACGCTGTGC----- 1108
 QY 1615 tctctgaagaggtccctctctcttttgaactctctcaagccccctccctctctgctgag 1674
 Db 1108 ---CTCAAGGAGGAGCAGCGCGCTACGACCTGTGCGCGCGCTCCGAGAACTCGCA 1163
 QY 1675 ggaactccataaagtgaatacaaaagagggagtgatgagtgatgttcttcagtgacca 1734
 Db 1164 GGACGCGCTCGCGTGTGTCAGCACAGCGGGGAGCAGATGAAGGTGTACAAGTGCGRACA 1223
 QY 1735 ctgcccagct 1794
 Db 1224 CTGCGGGTGTCTTCTTGGATCAGCTCATGTATACCACTCCACATGGGCTGCCAGCGCTT 1283
 QY 1795 ccgtgatccctttagtgtaacatgtgtggtatcgagccacacacacacacacacacacacacac 1854
 Db 1284 CCGTGATCTTTGAGTCAACATGTGCGGTACCCACAGCCAGGACCGGTACGAGTTCTC 1343
 QY 1855 ctctcaatcgccagagagagacag 1881
 Db 1344 GTCCGACATAACGCGAGGGGAGCACC 1370

RESULT 5

V66969
 ID V66969 standard; cDNA; 1386 BP.
 AC V66969;
 DE 14-JAN-1999 (first entry)
 DT Human Ikaros encoding cDNA.
 KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
 KW differentiation marker; immune system; corpus striatum; AIDS;
 KW Alzheimer's disease; ss.
 OS Homo sapiens.
 FH Key
 FT CDS
 FT Location/Qualifiers
 1..1386
 /product= "Ikaros"
 US5824770-A.
 PD 20-OCT-1998.
 PF 05-JUN-1995; 465590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 98-582621/49.
 DR P-PSDB: W72672.
 PI Ikaros poly:peptide(s) - useful for treating disorders of immune
 PT system or corpus striatum
 PS Disclosure: Column 55-58; 111pp; English.
 CC The present invention describes a purified peptide having at least one
 CC of the following properties: (a) it stimulates transcription of a DNA
 CC sequence under the control of a delta A element, an NFkB element or an
 CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
 CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
 CC consensus sequence; (c) it competitively inhibits the binding of a
 CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
 CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
 CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
 CC (e) it inhibits protein-protein interactions of transcriptional complexes
 CC formed with naturally occurring Ikaros isoforms. The proteins, provided
 CC that they stimulate gene transcription under the control of delta A
 CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

CC competitively inhibit binding of naturally occurring Ikaros isoforms to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
 CC inhibit protein-protein interactions of transcriptional complexes with
 CC naturally occurring Ikaros isoforms, can be used to treat immune system
 CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
 CC Alzheimer's disease. The present sequence encodes a specifically
 CC claimed human Ikaros protein.
 SQ Sequence 1386 BP; 350 A; 403 C; 395 G; 238 T;

Query Match 20.8%; Score 411.8; DB 1; Length 1386;
 Best Local Similarity 63.3%; Pred. No. 5.3e-94;
 Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;
 QY 700 cagcagacacacacagtgaggatgaactcgacgtgtgcgggttatccctgcattagctt 759
 Db 162 CATTTCGACTTCTTAACGGAATAAAGTGTATATCTGTGGATCATTTTCATCGCGGCC 221
 QY 760 caacgtctttagttcataagcgaagccataccggcgaaccccgcttccagtgtaataca 819
 Db 222 CAATGTGCTCTATGGTTTCACAAAAGAACCCACTGTGGAGAAGCGCCCTTCCAGTGCATCA 281
 QY 820 gtgcggggcatcttttactcagaaagttaacctcctcctcctcattataaactgcacaggg 879
 Db 282 GTGCGGGGCTCATATCACCAAGAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGGG 341
 QY 880 ggaataaaccttttaagtgtcacctctgcaactcgcacgacgacgaagagagatgcgtcac 939
 Db 342 GGAGAAGCCCTTCAATGCCACCTCTGCAACTAGCCCTGCCGCGGAGGAGCGCCCTCAC 401
 QY 940 gggacaccttaggacacattctgtggagagccgtacaaagtgtgagttctgcggaaagaag 999
 Db 402 TGGCCACCTGAGGACGACCTCCGTGTGTAACCTCACAATGTGGATATTGTGCCCGAAG 461
 QY 1000 ctacaagcagagaagctccctgagagagcagaagcagcgtgcgagcttttcttcagaa 1059
 Db 462 CTATAACAGCGAAGCTCTTTAGAGGAACATAAAGAGCGCTGCCAACACTACTTGGAAAG 521
 QY 1060 ccttgacct--gggggacgctgcaagtgtgaggcgaagacacacatacgaagc----- 1108
 Db 522 CATGGGCTTCCGGGACACACTGTACCCAGTCAATTAAGAGAGAACTAAGCACAGTGAAT 581
 QY 1108 -----cgagatggaaagtggagagctctcgtctcgtgacagattagcaag 1152
 Db 582 GGCAGAAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTGCTGTGCACAGACTAGCAAG 641
 QY 1153 caatgtgctaaagaaaaagctcgatccctcagaaattcattcgtgtagaagggcactg 1212
 Db 642 TAATGTCCCAAAAGCTAAGAGCTCTATGCCCTCAGAAATTTCTTGGGGACAAAGGCCCTGTC 701
 QY 1213 ctctgatgccaaactacatcccggtctacatgtacgagaaggaagacgatgatgcagac 1272
 Db 702 CGACACCCCTTACGACAGTGC-----CAGTACGAGAAGGAGAACCAATGATGAAGTC 755
 QY 1273 ccgagatgatggaccaagccatcaataacgcatcagctatctagggggtgaagccttcg 1332
 Db 756 CCACGTGATGGACAAGCCATCAACAACGCCATCACTACTGCGGCGCGAGTCCCTCGC 815
 QY 1333 ccccttagtcagacacgctcctccacctctgagatggtccagtcattcagcagtgt 1392
 Db 816 CCCCTGGTGCAGACGCCCCCGG---CGGTTCGAGAGTGTTCCTCCGCTCATACAGCCGAT 872
 QY 1393 gtaccccata-----gcaacttaactcgggccgatgacgaatgggggccccgca 1440
 Db 873 GTACCACTGCACAGGCGCTCGGAGGCGACCCCGCTCCCAACCACTCGGCCCCAGGACAG 932
 QY 1441 ggaagatgaaagaaacggtatcctctccagagaagattcttctcttgaagagtgct 1500
 Db 933 CGCGGTGGATACCTGTGCTGTCTCTCAAGGGCAAGTTGTGTGCCCTCGGAGCGCAGGC 992
 QY 1501 gtcccccataaacagtgccagagactccacagacacagacagacacacacagagatc---- 1557


```

Db 1095 GTACAGCTGCACAGGCGCTCGGAGGACACCCCGCTCCACACCTCGGCGCCAGGACAG 1154
Qy 1441 ggagatgaaagaaacggatccctcctgagagaaagatcttcttctgaacaggtct 1500
Db 1155 CGCGGTGGAGTACCTGCTGCTCTCTCAAGGCCAAGTGTGTGCTCGGAGCGCGAGGC 1214
Qy 1501 gtcccccataaagtgccaggagctccacagacacacagacacacacagagatc---- 1557
Db 1215 GTCCCGGAGCAACAGCTGCCAAGACTCCACGACACCGAGAGCAACACGAGGAGCAGCG 1274
Qy 1557 --gcaacatctaccagcaaacgagtggtctctcccccagcccgcaatggatgcc 1614
Db 1275 CAGGGCTTATCTACTGACCAACACATCGCCGAGCGCGCAACGCTGTCTG----- 1330
Qy 1615 tcttctgaagagtgctcctgctcttttgaactctcctcaagcccccctccactctgcctgag 1674
Db 1330 ---CTCAAGGAGGAGCAGCGCGCTACGACCTCTGTGCGCGCGCCCTCCGAGACTCGCA 1385
Qy 1675 ggaactcaataaagtatcaacaaagagggaggtgatggatgtgttcttctgagtgacaa 1734
Db 1386 GGACGCGCTTCGCGGTGTGTCAGACACGAGCGGGAGCAGATGAAGTGTACAAAGTGCGAACA 1445
Qy 1735 ctgcccagctctctctctagattatgtgttccacatccacatgggtggtccatggtt 1794
Db 1446 CTGCGGGTGTCTCTCTGATACAGTCAATACCATCTCCATGCTGCTGCTGCTGCTGCTGCT 1505
Qy 1795 ccgtgatccctttgagtgtaacatgtgtgctatcgaagccacagatcgatgagttctc 1854
Db 1506 CCGTGATCTTTTCAGTGCACCATGTGCGGTACCACAGCCAGCAGCGGTACGAGTTCTC 1565
Qy 1855 ctctcaatcgcagagagagacacagaccatgttgaaagtgcac 1899
Db 1566 GTCGCACATAACGCGAGGGAGCAGCCGCTTCCACATGACGTAAGC 1610

```

RESULT 4

```

T16060
ID T16060 standard; cDNA; 1386 BP.
AC T16060;
UT 09-MAY-1996 (first entry)
DE Human Ikaros cDNA h1k-1.
KW Ikaros; transgene; transgenic animal; transgenic mouse; h1k-1;
KW immunocompromised; immune system disorder; nervous system disorder;
KW animal model; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT exon 1..255
FT FT /*tag= a
FT FT /label= Exon-3
FT FT 256..423
FT FT /*tag= b
FT FT /label= Exon-4
FT FT 424..549
FT FT /*tag= c
FT FT /label= Exon-5
FT FT 550..681
FT FT /*tag= d
FT FT /label= Exon-6
FT FT 682..1383
FT FT /*tag= e
FT FT /label= Exon-7
FN W09604372-A1.
PD 15-FEB-1996.
PF 28-JUL-1995; U09345.
PR 29-JUL-1994; US-283300.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 96-129389/13.
DR P-PSDB; R92015.
DR Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
PT severely immuno-compromised and can be used as model to determine
PT effects of treatment for immune and nervous system disorders

```

```

PS Disclosure: Fig 2: 102pp: English.
CC An almost full-length cDNA sequence (T16060) codes for part
CC (R92015) of the human Ikaros protein, a master regulator of
CC haematopoietic differentiation. It was isolated from a T-cell
CC line Jurkat cDNA library using a mouse Ikaros cDNA clone as probe.
CC The human Ikaros gene maps to pl1.2-pl3 on chromosome 7.
CC Different isoforms (see R92014 and R92016-19) of mouse Ikaros have
CC also been isolated. Transgenic animals, pref. mice, having a
CC mutated Ikaros transgene, esp. a mutation that alters the DNA
CC binding domain of the Ikaros protein, are used as models to
CC determine the effects of treatments for immune or nervous system
CC disorders.
CC Sequence 1386 BP; 350 A; 403 C; 395 G; 238 T;
SQ

```

Query Match 20.8%; Score 411.8; DB 1; Length 1386;

Best Local Similarity 63.3%; Pred. No. 5.3e-94;

Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

```

Qy 700 cagcagacacacacagtggaagatgaactgcgacgtgtgctgggttatcctgcattagctt 759
Db 162 CATTTCGACTCTCTAACGGAAACTAAAGTGTGATATCTGTGGATCATTTGCATCGGGCC 221
Qy 760 caacgtcttgatgttcataagcgaagccataccgagcaacgcccttccagtgtaatca 819
Db 222 CAATGTGCTCATGGTTACAAAAGAACACACTGGAGAACGGCCCTTCAGTGCAATCA 281
Qy 820 gtgcgggggcatctttactcagaaggttaacctccctccgtcatattaaactgcacacggg 879
Db 282 GTCGGGGCTCATTCACCAGAGGGCACTGCTCCGGCAGCATCAAGCTGCTATTCGG 341
Qy 880 ggaataacacctttaaagtgtcaactctgcaactacgcgcacatcccaagagagatgcgtcac 939
Db 342 GGAGAAAGCCCTTCAAAATGCCACTCTGCAACTACGCTGCGCGGAGGAGGACGCCCTCAC 401
Qy 940 gggacaccttagcacacattctgtggaagcgtacaaagtgtgagttcttcggaagaag 999
Db 402 TGCCACCTGAGGAGCAGCTCCCTTGTAACTCAAAATGTGGATATTGTGCGCGAAG 461
Qy 1000 ctacaagcagagaagctccctgagagagcacaaggaagcgtcgcgagcttttcttcagaa 1059
Db 462 CTATAACACGGAACGCTTTAGAGGAACATAAAGAGCGCTGCCACACACTACTTGGAAAG 521
Qy 1060 ccttgacct--ggggagcgtgcaagtgtgagggaagacacacatcaaac----- 1108
Db 522 CATGGGCTTCCGGGACACACTGTACCCAGTCTATTAAGAGAACTAAAGCACAGTGAAT 581
Qy 1108 -----cgagatgggaagtgaagagcgtcgtcctcctgcagagattagcaag 1152
Db 582 GGCAGAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTCTGGACAGATAGCAAG 641
Qy 1153 caatgtggcttaagcgaagagctcgatgcctcagaattctcgtgagagcggaagcactg 1212
Db 642 TAATGTGCCCAACGTAAGAGCTCTATGCTCAGAAATTTCTTGGGACAAAGGCGCTGTC 701
Qy 1213 cttcagatgccaactacaatccccggctacatgtacgaggaagagagatgatgcagac 1272
Db 702 CGACAGCGCCCTACGACAGTGC-----CAGGTACGAGAGAGAGAGAAATGATGAAGTC 755
Qy 1273 ccggatgtagcacaagccatcaataacccatcagctatctagggctgaagcccttcg 1332
Db 756 CCACGTGATGGACCAAGCCATCAACACGCCATCACTACTCTGGGGCGGAGTCCCTGCG 815
Qy 1333 ccccttagtccagactccgctcctccactctgcagagtggtccagatcatcagcagtg 1392
Db 816 CCGCTGGTGCAGACGCCCCCGG---GCGGTTCGAGAGTGGTCCCGGTCTATCAGCCGAT 872
Qy 1393 gtaccocata-----gcaettactcgggcgcatatgccaatggggggcccgca 1440
Db 873 GTACAGCTGCACAGGCGCTCGGAGGGGACCCCGCGCTCCAACTCTCGGCCCGGAGACAG 932
Qy 1441 ggagatggaagaaacggatcctcctgccagagaagatcttgccttctgaaagaggtct 1500

```

CC committed lymphoid progenitors and in T and B cells, and is a
 CC transcriptional activator of a lymphoid gene. Aiolos cDNA can be
 CC used to produce recombinant Aiolos polypeptide (W15575) in host
 CC cells. The Aiolos polypeptide, coding sequence or cells expressing
 CC Aiolos may be used to treat a disorder in an animal, esp. by gene
 CC therapy. Such disorders include T-cell leukaemias and lymphomas.
 CC Non-wildtype gene structure or expression is indicative of a risk
 CC for such a disorder. Transgenic animals with an Aiolos transgene
 CC are provided.
 SQ Sequence 628 BP; 207 A; 128 C; 155 G; 138 T;

Query Match 25.0%; Score 496.8; DB 1; Length 628;
 Best Local Similarity 88.1%; Pred. No. 2.1e-115;
 Matches 553; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

Qy 569 gacagagatgagaacattatgaagcggagcccatggagatgcagagagagtgaaatg 628
 Db 1 GAAGAGATGAGAAATGTTTAAAGTCAGAACCCATGGGAAATGCAGAGAGCCTGAATC 60

Qy 629 cettacagctatgcaagagaatacacagcgactatgaagcattaaagctggagagaca---c 685
 Db 61 CCTTACAGCTATTCAAGAGATAATATGAATATGAAACATTAAAGTTGGAGAGACATGTT 120

Qy 686 gtgcctatgacaacagcagacacacacacacacacacacacacacacacacacacac 745
 Db 121 GTCTCATTCGATAGTAGAGAGCCACACAGTGGAAAGATGAATCGGATGTGTGTGGATTA 180

Qy 746 tctgcattagtttcaaacctcttgatggttcataagcgaagccatccacggcggaagcccg 805
 Db 181 TCTTCATCAGCTTCAATGCTTATGTTTATAGCGAGAGCCATCTGTTGAAGCCCA 240

Qy 806 ttccagtgtaactagtgccgggcatcttttactcagaaggttaacctctcctcgctcatt 865
 Db 241 TTCCAGTGTAAATCAGTGTGGGCATCTTTTACTCAGAAAGGTAACCTCTCCGCCACATT 300

Qy 866 aaactgcacaggggggaaaaaaccttttaagtgtcacctctgcaactaagcagcatgccaaagg 925
 Db 301 AAATGTCACACAGGGGAAACCTTTTAAAGTGTCACTCTGCAATATGTCATGCCAAAGA 360

Qy 926 agagatgcgctcacgggacacaccttaggacacattctgtggaagccgtcacaagtgtgag 985
 Db 361 AGAGATGCGCTCACGGGCATCTTAGGACACATTCTGTGGAGAAACCTTACAAATGTGAG 420

Qy 986 ttctcgggaagctacacagcagagaagctccctcggagagcacaaggaacgctgccga 1045
 Db 421 TTTTGTGGAAGGAGTTTACAAGCAGAGAAGTTCCTTGTAGGAGCACAAGGAGCGCTGCCGT 480

Qy 1046 gcttttttcagaacctgacctggggagcgtgcaagtgtgagggcgaagacacacataaa 1105
 Db 481 ACATTTCTTTCAGAGCACTGACCACGGGGACCTGCAAGTGGCGAGGCAAGACACATCAAA 540

Qy 1106 gccgagatgggaagtgaagagactctcctcgtgagacagattagcaagcaatgtgctaa 1165
 Db 541 GCAGAGATGGGAAGTGAAGAGACTCTCGTACTTGGACAGATTAGCAGCAATGTGGCAAAA 600

Qy 1166 cgaaaaagctcagctcctcagaaattca 1193
 Db 601 CGAAAAAGCTCAATGCTCAGAAATTCA 628

RESULT 3
 ID Q44980
 AC Q44980;
 DT 21-OCT-1994 (first entry)
 DE Human Ikaros peptide coding sequence.
 KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
 OS corpus striatum; regulatory gene; ss.
 FH Key
 FT cds
 FT Location/Qualifiers
 1..1611
 /*tag= a

FT W09406814-A.
 FN 31-MAR-1994.
 PD 14-SEP-1993; U08743.
 PR 14-SEP-1992; US-946233.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 94-118387/14.
 DR P-PSDB; R46964.
 FT T-cell pathway regulatory gene, Ikaros - encodes family of unique
 FT zinc finger proteins, useful for treating immune system disorders
 PS Claim 13; Page 44-46; 112pp; English.
 CC The Ikaros gene encodes a zinc finger protein which can be used in a
 CC therapeutic composition to treat animals with an immune system
 CC disorder. It may also be used for assessing whether a subject is at
 CC risk for an immune disorder. It is of particular use in treating a
 CC disorder of the corpus striatum.
 CC Sequence 1611 BP; 375 A; 480 C; 272 T;

Query Match 20.9%; Score 413.8; DB 1; Length 1611;
 Best Local Similarity 63.1%; Pred. No. 1.8e-94;
 Matches 785; Conservative 0; Mismatches 397; Indels 63; Gaps 7;

Qy 700 cagcagac 759
 Db 384 CATTCGACTTCTTAAACGGAACCTAAAGTGTGATATCTGTGGATCATTTTCATCGGGCC 443

Qy 760 caacgtctgtgttcttaagcgaacccataccggaacccgctccagtgtaatca 819
 Db 444 CAATGTGCTCATGTTTACAAAAGAGCCACCTGGAGACGGCCCTTCCAGTGCAATCA 503

Qy 820 gtgcgggac 879
 Db 504 GTGCGGGCCCTCATTCACCCAGAGGGCAACCTCTCCGGCACATCAAGCTGCATTCGG 563

Qy 880 ggaac 939
 Db 564 GGAGAAGCCCTTCAATGTCACCTCTGCACTAGCCCTGCGCGGAGGGAGCGCCCTCAC 623

Qy 940 gggac 999
 Db 624 TGGCCACTCTGAGGAGGACACTCCGTGTGTAACCTTCAAAATGTGATATTGTGCCCAAG 683

Qy 1000 ctacaagcagagaagctccctcggagagcacaaggaacgctgcccagagcttttcttcagaa 1059
 Db 684 CTATAACAGCGAAGCTTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAG 743

Qy 1060 ccttgacct--ggggagcgtgcaagtgtgaggaagacacacacacacacacacacac 1108
 Db 744 CATGGGCTTCCGGGCACACTGTACCCAGTCTATTAAAGAGAAACATAAGCACAGTGAAT 803

Qy 1108 -----cagagatgggaagtgaagagctctcctcgtgagacacacacacacacacac 1152
 Db 804 GGCAGAGACACTGTGCAAGATAGGATCAGAGAGATCTCTCGTCTGGACACACTAGCAAG 863

Qy 1153 caatgtgctgaagcgaagaaagctcgtgcctcagaaattcatcgtgagagagcgcactg 1212
 Db 864 TAATGTCCCAACGTAAGAGCTCTATGCTCTCAGAAATTTCTTGGGACACAGGGCCCTGC 923

Qy 1213 cttcgatgcgaactac 1272
 Db 924 CGACACGCCCCACGACAGTGC-----CACGTAGGAGAGAGAGAGAAATGATGAAGTC 977

Qy 1273 ccggaatgagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaag 1332
 Db 978 CCAGTGATGACCAAGCCATCAACAAAGCCATCAACTACCTGGGGCCGAGTCCCTGCG 1037

Qy 1333 ccccttagtcagactcgcctcgtccacacctgagatggtccagtcacacagcagtg 1392
 Db 1038 CCCGCTGTGTCAGACGCCCCCGGG---CGGTTCGAGGTGTCCCGGTCTCATCAGCCGAT 1094

Qy 1393 gtaccccaata-----gcacttaactcggggccgatatgccaatggggggcccccga 1440

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 1999, 21:32:29 ; Search time 70.34 Seconds
(without alignments)
7056.882 Million cell updates/sec

Title: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagagcgacacgcgtctcg.....gaactcaaacccacctcgag 1984

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1980.8	99.8	1984	1	Mouse Aiolos cDNA.
2	496.8	25.0	628	1	Human Aiolos parti
3	413.8	20.9	1611	1	Human Ikaros pecti
4	411.8	20.8	1386	1	Human Ikaros cDNA
5	411.8	20.8	1386	1	Human Ikaros encod
6	411.8	20.8	1386	1	Human Ikaros isofo
7	411.8	20.8	1551	1	Human Ikaros isofo
8	379.6	19.1	2049	1	Murine Ikaros cDNA
9	379.6	19.1	2049	1	Mouse Ikaros encod
10	379.6	19.1	2049	1	Mouse Ikaros isofo
11	348	17.5	1788	1	Mouse Ikaros isofo
12	346.4	17.5	1788	1	Murine Ikaros gene
13	346.4	17.5	1788	1	Murine Ikaros cDNA
14	346.4	17.5	1788	1	Murine Ikaros enco
15	269.6	13.6	1004	1	Ikaros cDNA. Trans
16	269.6	13.6	1004	1	Ikaros isoform enc
17	269.6	13.6	1004	1	Human Ikaros cDNA.
18	236.8	11.9	1170	1	Mouse Ikaros isofo
19	235.2	11.9	1170	1	Murine Ikaros cDNA
20	235.2	11.9	1170	1	Mouse Ikaros encod
21	200.4	10.1	714	1	Ikaros protein enc
22	191.2	9.6	1296	1	Murine Ikaros cDNA
23	191.2	9.6	1128	1	Murine Ikaros cDNA
24	191.2	9.6	708	1	Ikaros protein enc
25	191.2	9.6	1128	1	Mouse Ikaros encod
26	191.2	9.6	1296	1	Mouse Ikaros encod
27	191.2	9.6	1296	1	Mouse Ikaros isofo
28	191.2	9.6	1128	1	Mouse Ikaros isofo
29	109.4	5.5	168	1	Ikaros protein enc
30	104.6	5.3	168	1	Ikaros protein enc
31	73	3.7	1663	1	Renal cancer assoc
32	70.8	3.6	1683	1	Human SRE-2BP anal
33	70.8	3.6	2168	1	Human SRE-2BP anal
34	70	3.5	353	1	EST clone AW95. Ne
35	66.4	3.3	264	1	DNA encoding a nuc
36	66.4	3.3	264	1	DNA encoding an an
37	66.4	3.3	264	1	Zinc finger protei
38	66.2	3.3	2680	1	Myc-binding zinc-f
39	64.8	3.3	1476	1	Zinc finger protei
40	64.4	3.2	323	1	Human gene signatu
41	63	3.2	2133	1	Human OTK18 gene.
42	63	3.2	3754	1	Human OTK18 gene.
43	62.8	3.2	270	1	EST clone CO1069.

44 61.4 3.1 3776 1 T42903 TRP-1 protein codi
45 60.2 3.0 3720 1 Q80513 Genetic locus bcl-

ALIGNMENTS

RESULT 1
T60490
ID T60490 standard; cDNA; 1984 BP.
AC T60490;
DT 07-JUL-1997 (first entry)
DE Mouse Aiolos cDNA.
KW Aiolos; transcription activator; immune system; T lymphocyte;
KW B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW transgenic animal; ss.
OS Mus sp.
FH Key
FT cds
Location/Qualifiers
374..1897
/*tag= a

PN M0371471A-A1.
PD 24-APR-1997.
PF 17-OCT-1996; U16774.
PR 18-OCT-1995; US-005529.
PR 14-MAY-1996; US-017646.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K, Morgan BA;
DR WPI; 97-245047/22.
DR P-PSDB; W15574.
PT Aiolos polypeptide and corresponding DNA - used to reconstitute a
mammalian immune system, for the treatment of T cell leukaemia(s)
PT and lymphoma(s)
PS Disclosure; Page 73-76; 115pp; English.
CC A cDNA clone (T60490) corresponds to the mouse Aiolos gene, a
homologue of Ikaros whose expression is restricted to lymphoid
lineage. Aiolos cDNA was isolated from a mouse spleen cDNA library
using a probe that spanned nucleotides 796-1156 of the sequence.
CC Aiolos (W15574) can form dimers with Aiolos or Ikaros polypeptides,
is expressed in committed lymphoid progenitors and in T and B cells,
and is a transcriptional activator of a lymphoid gene. Primers
CC (see also T60494-95) based on mouse Aiolos gene exons were used to
CC amplify a human partial Aiolos cDNA (see also T60491). Aiolos cDNA
CC can be used to produce recombinant Aiolos in transformed host cells.
CC The Aiolos polypeptide, coding sequence or cells expressing Aiolos
CC may be used to treat a disorder in an animal, esp. by gene therapy.
CC Such disorders include T-cell leukaemias and lymphomas. Transgenic
CC animals are provided with an Aiolos transgene.
SQ Sequence 1984 BP; 512 A; 570 C; 535 G; 367 T;

Query Match 99.8%; Score 1980.8; DB 1; Length 1984;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 cagagcgacacgcgtctcttcgtgacacgcctcattcccggtgtttctcaag 60
Db 1 CACGAGCGCACGCGTCCGGCTCTCTTGGACACGCGCTCATCCCGGTGTTTCTCAAG 60
Qy 61 tagacgtcccgagacggtctgagcactgtttccacgcatcagggttcctcagcgtt 120
Db 61 TAGACGTCCTCCGAGACGGTCTGCTGAGGCACGTGTTTCCACGCGATCAGGGTTCCTCAGGCTT 120
Qy 121 gacattcaaaagtgggtgcggaaaccccgcgacactcgagcggtgtctttaaagcgccgccca 180
Db 121 GACATTCAAAAGTGGGTGCGGAACCCGCGCACTCGAGGCGTGTCTTTAAAGCGCGCCCA 180
Qy 181 gccagcgccgctctaaactcgcccccggtgtcgcggcggtcccgctcatctgccc 240
Db 181 GCCAGCGCCGCTCTAACTCGCGCGCGGTGCGGGGCTCCCGGCTCATCTGCGCC 240
Qy 241 gacgacacgagcgtatcccgggacctctcccgccggaatctcccgccgacgcgggg 300
Db 241 GACGCGACGAGGATCCCCGGGGCTCTCCCTGCGCCCGGAATCTCCCGCCGACCGCGCGG 300

YKORSLEERHKKRCHNYLONVMEAGGIVHHASPMDECKNPETMMNNNSLMPFRP
AVTERLASNKKRSKSTPPQRFELGKLMRYGYPDLHFDMAEYKEAEELIQSOMDQAINN
AITYLGDALURPLIHSAAMPEVPIVSSLYSQVYHPARVPTRETSRSDSNNDNMND
GPISLIRPNHPQERASPSNCLDITDESSEHAPAILPFGNHALNSKRKQSSVYPKE
HGKPLDATKSSLSGSKDVYRVFNSEGEQIRAFKCEHRCRVFLDHNVTIHMG"

BASE COUNT 295 a 233 c 244 g 214 t
ORIGIN

Query Match 7.4%; Score 146.6; DB 4; Length 986;
Best Local Similarity 52.0%; Pred. No. 2.1e-23;
Matches 508; Conservative 0; Mismatches 394; Indels 75; Gaps 5;

Qy	878	ggggaaaaacccctttaagtgtcacctctgcaactacatgcagccaaaggagagatgcctc	937
Db	12	GGGGAGAGCGGTTTAAATCTTCATCTTCAGCTATGCATGTCGGAGAGAGACGCACTT	71
Qy	938	acgggacaccccttagacacattctgtgagagccgtacaaagtgtgattctcggaaga	997
Db	72	ACTGCCACCTCAACACACTCCGTTGGCAACCTCACAAGTCAACTACTGTGGTCGA	131
Qy	998	acctaaagcagagaagctccctggaggagcacaagaaagcgtccgagctttcttcag	1057
Db	132	AGCTACAAGCAGAGAGGCTCTTTGGAGGACACAAAGGAGAGGTGTCACAACTACTTCAA	191
Qy	1058	aacctgacctgggggagcgtgaagtgtg-----	1089
Db	192	AATTTGGGTGGAGGCTGGCCAAAGGATTTGTGCACCATGCATCTCCTATGGACGAGTGT	251
Qy	1089	-----aggaagcacatcaaacgagagatgggaagtggagagagctctc	1132
Db	252	AAGAAATCCAGAGACCATGATGGAACACACATGCTCTAATGCCATTTGAGAGACCGCT	311
Qy	1133	gtcctgacagattagcaagcaattgtggttaagcgaagaaagctcgatgcctcagaaattc	1192
Db	312	GTTATAGAGACTTGCAGCAACATGGGAAGCGTAAGAGCTCCACCCACAGAGGTTT	371
Qy	1193	atcgtgagaagcggcactgcttg---atccaaactacaatcccggtacatgaag	1249
Db	372	TTGGGTGAGAAGCTCATGAGATATGGATATCCTGACTTGCACTTTGATATGGCCATATGAG	431
Qy	1250	aaggagacgagatgacgagaccgagatgatgacaaagccatcaataagcccatcagc	1309
Db	432	AAAGGCGAGAAATATCCAGTCAAAATGATGATCAGGCGATTACAAATGCTATAACA	491
Qy	1310	tatctaggggtgaagccttcgcccccttagtcagactccgctgtctccacactctgag	1369
Db	492	TACCTGGGAGCAGATGCATCGGCTCTCATACCACTCAGCAGCTGCCATGCCCTGAG	551
Qy	1370	atggtcccgatcagcagtggttaccccaatagcaacttaactcgggc-----	1418
Db	552	GTTCACCAATTTGTCAGTTCACTTTACTCTCAGGTCTATCATCCAGCCAGAGTGGAGAGA	611
Qy	1418	-----gatagccaaatggggcccgagagatggaaagaaacggatcctc	1465
Db	612	CCCACGACGGGAGACATCAGACGACACGATACACACATGGATGGCCCAATCTCCCTT	671
Qy	1466	ctgccagagaagattgtcccttctgaacgaggtgtgtcccccataaacagtgcccagga	1525
Db	672	ATCAGACCAAGAACACCCTCAGGAGAGGGAGCATCACCCAGCAACAGCTGCTTGAT	731
Qy	1526	tccacagacaccgacagcaaacacgagagatggccaacatctctaccaggaagccacgtg	1585
Db	732	ACTACTGACTCAGAAAGACGACCATGAG--CACCGGCAATCCTACCGGCAATCATGCCT	789
Qy	1586	gtcctccccagcccgcaatgggtgctctctgaaggaggtccctcgctcttttga	1645
Db	790	TAAATTCAAAGAGAAAGCAAGTTGGGTGATCCCAAGGAACATGGCAAC-CATTGGAT	848
Qy	1646	ctcctcaagcccccctcccatcgtcctgagggactcccatcaaaagtgatcaacaagaagg	1705
Db	849	GCCACAAAGTCCACGCTTGGGTTCTCCCAAGGATGCTACAGAGTTTTTAATAGTGAAGGA	908

Qy	1706	gaggtgatggatgtgttttgatgtgaccactgcccagtcctcttctctagattatgtgatg	1765
Db	909	GAGCAGATTAGGCGCATTTAAATGTGAGCACTGTCTGGGTTCTGTCTTGACCATGTTATG	968
Qy	1766	ttcaccatccacatggg	1782
Db	969	TATACGATCCATATGGG	985

Search completed: November 6, 1999, 06:39:34
Job time: 4688 sec

Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
Location/Qualifiers
1. .2079
/organism="Oncorhynchus mykiss"
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
81..1232
/genes="Ikaros"
81..1232
/genes="Ikaros"
/notes="Ik-8 isoform; similar to mouse and human Ikaros/LyP-1; alternatively spliced form missing exons 3 and 6"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53433.1"
/db_xref="PID:g2062738"
/db_xref="GI:2062738"
/translation="MEMERAQMSQMPGRDPPNDVSENDAMPIDPDLSSNNLQ HNRGKELGEPFOCTQGASFTQGNLLRHILKHSGEKFKCHLCNYACRRRDAL SGLRTHSVKPKHCAYGRSYKORSLEEKHERCHNYLQCMGLQNSIYTGKRFNSI SFGGPELMQPHVIDOAINSAINYLGAESLRLIOTSPSTDGMVMSMYPLHRPPA EGHLSAKSAENALLLAKSASKEKSPSHGSDSTDTESNEEKAGVAGSGLI YLTHITSGVRNGLPLVKEOQRYEAMRASIEIASEGFKVLSGEQVRYRCEHC RILFLDHVWYTHMGCHGRDPRFECNLCGRSQDRFESSHMRGHEHRI"
BASE COUNT 593 a 500 c 464 g 521 t 1 others
ORIGIN

Query Match 8.3%; Score 164; DB 4; Length 2079;
Best Local Similarity 52.4%; Pred. No. 2.6e-27;
Matches 579; Conservative 0; Mismatches 380; Indels 147; Gaps 4;

QY 794 ggcgaacgcccgttcagtgtaacatcagtcgagggcattcttttactcagaaggttaacctc 853
DB 243 GGAGAGGTCATTCAGTGCACCCAGTGGCGCTTCTTTACCCAGGAAGGCAACCTG 302

QY 854 ctccgtcaatataaactgcacacgaggggaaaaaaccttttaagtgtcaacctctgcaactac 913
DB 303 CTACGTACATCAAGTCCACTCAGGAGAGAAGCCCTTCAAGTGTACCTGTGCAACTAT 362

QY 914 qcatgccaaagagagatcgctcagggacactcttaggacacatctctgtggagaagccg 973
DB 363 GCCTGTCCGCGAGAGACGCCCTCAGCGGTACCTGGTACCACCTCTGTGGAAACCC 422

QY 974 tacaagtgtgagttcttcggaagaagtacaaagcagagaagctccctggagggacaaag 1033
DB 423 CACAAAGTGTCTTACTGTGGCGGAGCTACAGCAGCGTAGCTCTGGAGGAACACAG 482

QY 1034 gaacgtgacgagctttcttcagaacacctgacctggagggagacgtgcaagtgtggagca 1093
DB 483 GAGCGGTGTACAACTACCTCCAG----- 507

QY 1094 agacacataaagcagagatgggaagtggagagctctcgtcctggacagattagcaagc 1153
DB 507 ----- 507

QY 1154 aatgtgctaagcgaaaaagctcgatcgctcagaaaattcatcggtgagacgagcactgc 1213
DB 507 -----TGCATGGGGCTCCAGAACAGCATCTATACAGGTGAGAAACGC 548

QY 1214 ttgatgccaaactacaatcccgctacatgtacagagaagagaaacagagatgtgcagacc 1273
DB 549 TTCTCCAAACATCTCTTCGAGGGTGGCCGGGGGAAC-----TCATGCAGCCT 596

QY 1274 cggatgtgaccaaagcacaataaacgcatcagctatctagggcctgaagccttcgcgc 1333
DB 597 CATGTCTACCAACGAGGCATCAACAGTGCCTCACTACTCTGGTGGCCGAGTCACTCCGA 656

QY 1334 cccttagtccagactccgctgtctccacacctctgagatggtcccgatccatcagcagtg 1393
DB 657 CCGCTCATCAGACCTCCCAACCTCCTCTGACATGGGGGTGATGGGCTCCATGTACCC 716

QY 1394 taccocatagcacttactctgagggccgatatgccaatggggggcccgagagagatgaaaaag 1453
DB 717 CTCCAAGCCCTCTGCAGAGGGCCACGGCTGTGAGCAAGGAGACAGCCGCAAAAT 776

QY 1454 aaacgatctctccagagagaatcttgccttctgaacgaggtctgtcccccataac 1513
DB 777 CTTCTGTCTCGAAAGTCCAAGTCAGCTCCAGGAGAGACGGCTTCCCCAGCCAC 836

QY 1514 agtgcacagactccacagacacaccgacacacacacagga-----tcgc 1558
DB 837 AGTGGCCAGATTCCACCCACACGGAGAGCAACAATGAAGAAGGAGGTGTGCGGGCC 896

QY 1559 caacatctctacagcaaacacacgtgctctcccccagggcccaatggagtgctctt 1618
DB 897 TCAGGGCTCTACTACCTGACAAACACATCAGCTCCGGGTACGTAATGGCTGTGTC 956

QY 1619 ct-----gaagaggttcctcgtctttgaactcctcaagccccctccatctgcctg 1672
DB 957 CTGGTGAAGGAGGAGCAGCAGCGGAGTACGAGGCCATCGGGGCCAGTATCAGATTGCC 1016

QY 1673 agggactccatcaaatgatcaacaaagagggaggtgatgattgttttcgatgtgac 1732
DB 1017 TCGGAGGGTTTCAAGGTGCTGAGTGAGAGGGGAGCAGGTGAGGCATATCGCTGTGA 1076

QY 1733 cactgcacgtctctctctcttagattatgtatgttccaccatccacatggggtgcatg 1792
DB 1077 CACTGCCGATCTCTTCTTAGACCATGTGTATACACCATCCAGTGGGTGCGCATGCG 1136

QY 1793 ttccgtgatcccttgatgtaacatgtgtgctatcgaacgacgacgacgacgacgacgac 1852
DB 1137 TTCGAGAGCCCTTCGAGTGCAACCTCTGTGCCATTCGCGAGTACGACCGTACGAGTT 1196

QY 1853 tctctcactcgcagagagagca 1878,
DB 1197 TCGTCACATGACCCGAGGGGAGCA 1222

RESULT 15
AF024439
LOCUS AF024439 986 bp mRNA VRT 20-OCT-1997
DEFINITION Xenopus laevis ikaros-related transcription factor mRNA, partial cds
ACCESSION AF024439
NID 92547231
VERSION AF024439.1 GI:2547231
KEYWORDS African clawed frog.
SOURCE Xenopus laevis
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 986)
AUTHORS Turpen,J., Kelley,C., Mead,P. and Zon,L.
TITLE Bi-Potential Primitive-Definitive Hematopoietic Progenitors in the Vertebrate Embryo
JOURNAL Immunity (1997) In press
REFERENCE 2 (bases 1 to 986)
AUTHORS Turpen,J., Kelley,C., Mead,P. and Zon,L.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1997) Hematology, Children's Hospital, 300 Longwood Ave, Boston, MA 02115, USA
FEATURES
source location/Qualifiers
1. .986
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/codon_start=3
/product="Ikaros-related transcription factor"
/protein_id="AAB81280.1"
/db_xref="PID:g2547232"
/db_xref="GI:2547232"
/translation="IRLGEKPFKPCFSYACRRRDALTGHLRTHSVGPKHCNYCGRS

```

Query Match      9.4%; Score 187.2; DB 12; Length 684;
Best Local Similarity 65.1%; Pred. No. 1.5e-32;
Matches 336; Conservative 0; Mismatches 153; Indels 27; Gaps 3;

QY 700 cagcagaccacacagtggaagatgaactgcagctgagctgtgcgggttatctcattagctt 759
DB 168 CATTCGACTCTCTTACCGAAACTAAAGTGTATATCTGTGGGATCGTTTGCATCGGGCC 227
QY 760 caacgtcttgatggttcaagcgaagcattaccggaagcgcgcgttccagtgtaatca 819
DB 228 CAATGTGCTCATGTTTCACAAAGAGTCATACT---GAAAGGCGCTTTCAGTGCACCA 284
QY 820 gtgcggggcattctttactcagaaggttaacctctcctcgtcatattaaactgcacacggg 879
DB 285 GTGTGGGGCCTCTTACCCAGAAAGGCAACCTCTCGCGGCACATCAAGCTGCACCTCGG 344
QY 880 ggaataacaccttttaagtgtcaccctgtgaactacgcatgccaagagagagatgcgtcac 939
DB 345 TCAGAAGCCCTTCAATGCCATCTTTGCAACTATATGCTCGCGCGGAGGAGCGCCCTCAC 404
QY 940 gggacaccttaggacacattctgtgagaagccgtacaaagtgtgagttcttcggaagaag 999
DB 405 CGGCCACCTGAGGAGCGACTCCGTTGGTAAGCCTCACAAATGTGGATATTGTGCCGGAG 464
QY 1000 ctacaagcagagaagctccctggaggagcacaaggaagcgtgcgagcttttcttcagaa 1059
DB 465 CTATAACAGCGGAGAGCTCTTAGAGGAGCATAAAGAGCGGATGCCACAACACTACTTGGAAAG 524
QY 1060 ccttgacctggggg-----acgtgcgaagtgtggaggcaagacacataaagc----- 1108
DB 525 CATGGCCCTCGGGCATGTACCCAGTCTATTAAAGGAAGAACTAACCCACAGAGATGGC 584
QY 1108 -----cgaagtggaaagtgcagagctctcgtcctgcagacagattagcaagcaa 1155
DB 585 AGAAGACCTGTGCAGATAGAGAGAGAGAGGTCCCTTCTCTGGACAGGCTGGCAAGCAA 644
QY 1156 tgtggttaagcgaataaagctcgtgcctcagaaatt 1191
DB 645 TGTCGCCAAAGTAAGAGCTCTATGCCTCAGAAAT 680

RESULT 13
XLU92202      415 bp      mRNA      VRT      02-MAY-1997
LOCUS      Xenopus laevis/gillii Ikaros homolog (Ikaros) mRNA, partial cds.
DEFINITION
ACCESSION      U92202
NID      g2062743
VERSION      U92202.1 GI:2062743
KEYWORDS
SOURCE      Xenopus laevis/gillii.
ORGANISM      Xenopus laevis/gillii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
Xenopus.
REFERENCE      1 (bases 1 to 415)
AUTHORS      Hansen, J.D.
TITLE      Isolation and characterization of Ikaros homologues in the rainbow
JOURNAL      trout
REFERENCE      2 (bases 1 to 415)
AUTHORS      Hansen, J.D.
TITLE      Direct Submission
JOURNAL      Submitted (06-MAR-1997) Comparative Immunol., Basel CH-4005, Switzerland
IMMUNOL., Grenzacherstrasse 487, Basel
FEATURES
Source      Location/Qualifiers
1..415
/organism="Xenopus laevis/gillii"
/db_xref="taxon:8359"
/tissue_type="thymus"
/clone_lib="LG15.0"
<1..>415

/gene="Ikaros"
<1..>415
/genes="Ikaros"
/notes="similar to mouse and human Ikaros/LyF-1; composed
of exons 4, 5 and 6"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53435.1"
/db_xref="PID:g2062744"
/db_xref="GI:2062744"
/translation="GERPFCNOCGASFTQKGNLLRHILKHSGEKPKFCHMCNACRR
RDALSGHLRTHSVGPKGCGRSYKQBSLEEHKERCCHNYLQSLGLOHLYAVKEE
SKIDMAEDLSKIGSRSLVLDRLASNVAKRSMQP"
BASE COUNT      124 a      93 c      100 g      98 t
ORIGIN

Query Match      8.4%; Score 166.2; DB 4; Length 415;
Best Local Similarity 65.3%; Pred. No. 7.7e-28;
Matches 271; Conservative 0; Mismatches 123; Indels 21; Gaps 1;

QY 794 ggcgaacgcccgttccagtgtaatacagtcggtgggcatctttactcagaaggttaacctc 853
DB 1 GGTGAACGGCCTTTCAGTGCACACCACTGTGGTGCCTCTTACCCAGAGGGCAATCTC 60
QY 854 ctccgtcatattaaactgcacacgggggaaacaccttttaagtgtcacctctgcacatac 913
DB 61 CTTGCTCATATCAAAATACACTCTGTGTGAAAGGCCCTTAAATGTACATGTGTAACTAT 120
QY 914 gcatgccaaagagagatgcgtcacgcgggacacaccttttaggtgtcacctctgcacatac 973
DB 121 GCATGACGGCGCAGGGATGACCTCAGTGGTGCCTCAGGACACATCTGTAGGCAAACT 180
QY 974 tacaagtgtgagttcttcggaagaagctacaagcagaagctccctggaggagcacaag 1033
DB 181 CACAAGTGTGGTACTGTGGTGTAGTATAGCAACGAGAGCTCTTAGAGAGCATATAA 240
QY 1034 gaacgtgcgcgagcttttcttcagaacacctgacctgggggacgctgcaagtgtggaggca 1093
DB 241 GAGAGATGTCACAACTACCTACAAAGTATAGGCTTGCGGGCCATCTTTATGCTGTAAA 300
QY 1094 agacacatcaaa-----gccagatgggaagtgcagagagctctc 1132
DB 301 GAAGAGAGCAAAATGATATGGCAGAAAGACCTGTCCAAAGATTGGGTGAGAGATCACTT 360
QY 1133 gtcttgacagattagcaagcaatgtgcttaagcgaaaagctcgtgcctcaga 1187
DB 361 GTGCTTCACAGACTTGAAGTAACTGTCACCAAGCTTAAGAGCTCTATGCCTCAGA 415

RESULT 14
OMU92199      2079 bp      mRNA      VRT      02-MAY-1997
LOCUS      Oncochrynychus mykiss Ikaros homolog (Ikaros) mRNA, IX-8 isoform,
DEFINITION      complete cds.
ACCESSION      U92199
NID      g2062737
VERSION      U92199.1 GI:2062737
KEYWORDS
SOURCE      rainbow trout.
ORGANISM      Oncochrynychus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE      1 (bases 1 to 2079)
AUTHORS      Hansen, J.D.
TITLE      Isolation and characterization of Ikaros homologues in the rainbow
JOURNAL      trout
REFERENCE      2 (bases 1 to 2079)
AUTHORS      Hansen, J.D.
TITLE      Direct Submission
JOURNAL      Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for

```



```
Query Match      13.1%; Score 260.6; DB 4; Length 2309;
Best Local Similarity 56.0%; Pred. No. 4.7e-49;
Matches 718; Conservative 0; Mismatches 444; Indels 121; Gaps 7;

QY 700 cagcagaccacagtgaggagatgaactgcgagctgtgcgggttatctcctgcattagctt 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 CATCGCGCTCCCCAACGGGAAGCTGAAGTCGATATCTGTGGGATAGTTTGCATTGGGCC 489

QY 760 caacgttcttgatgttcaataagcgaagccatacc----- 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 CAATGTGTTGATGTTTCACAAAGCGAAGTCACACTGAAGAAGAAAGTTCAGTTTTTGAACA 549

QY 794 -----ggcgaagcccgctccagtgtaatacagtgccgggcatcttttactcagaagg 846
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 ACAAAGAGTGAAGGCCATTCCAGTGCATCAATCAATGTGGTCTCATCTACTCAGAAAGG 609

QY 847 taacctctcctgcgtcatattaaactgcacacggggggaacaccttttaagtgtcacctctg 906
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 TAACCTGCTCCGACACATCAAACTTCACTCTGGCGAGAAACCTTTCAAATGTCAACCTGTG 669

QY 907 caactacgcatgcaaaaggagagatgcgtctcagcgggacacctttagacacattctgtgga 966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 CAACATGCTTGGCGCGCAGAGACGCTCTCACTGGACATCTGCGCACTCATTCGGTTGG 729

QY 967 gaagcgtacaaagtgtgagttctcggaaagaagctacaagcagagaaagctccctgagga 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 AAGCCCCATTAAGTGTGCATTTGCGGACCGAGTTTCAAGCAGCGGAGCTCACTGGAGGA 789

QY 1027 gcacaaggaaacgtgcgcagctttcttcagaacctgaacctggggagcgtcgaagtgt 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 790 ACATAAGGAGAGATGTCACAACTACTTGCAGTGCATGGCGCTTCAGAACAA--CAATTAT 847

QY 1087 gagggcaagcacatcaaaagccagatgggaag-----tgagaga 1126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 ACAGTGAAGGAAGAGAACACGCCAGAATGAGCAGAGGAGGACATGCTGCACTGAGAGG 907

QY 1127 gctctgctctgacagattagcaagcaatgtgctgaagcgaagcaagctcgatgcctcag 1186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 GCCTTGGTGTACACAGGATAGTACAAATGTAGCTAAGCGTAAGAGCTCTATGCCACAG 967

QY 1187 aaattcatcggtgagaagcggcactgtctcgatgccaaactacaatccccggctacatgtac 1246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 968 AGGTTTGGGAGAGAATCGTGTGCAGAGCTATCTTTCGAGAGTGGCTCAGT----- 1022

QY 1247 gagaaggagaaacagatgatgcagaccgagatgatggaccgaacccaataacgcgaatc 1306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 -----GAGCTGATGTCAGGCCCATGTGATGATGATGATGATGATGATGATGATGATG 1069

QY 1307 agctatcagggtgaaagccttccgccccttagtccagactcgcgtctcccaactct 1366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1070 AGCTATCTGGGTGCAGAGTCTCTTGGGCGCTCTGTTTCAGACCTCTCTCTGGGTCCGCCGAC 1129

QY 1367 gagatgggtccagtcacagcagtggtgtaccccatagcaacttactcgggcccagatgcca 1426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1130 ATGCTGTGTCAGCCCTCTATCAACCTGCACAAGTCAACAACAGCTGAAGCAATGGCGTT 1189

QY 1427 atggggggcccgagagatggaagaaacgagatcctctcgcagagagatcttgccct 1486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1190 TCTGCTAAAGACAGCGCCGAGAGCACCTTCTCTTACTTCTTAAGTCCAAATCCGCTCT 1249

QY 1487 tctgaacgaggtctgtcccccataacagtgcccagactccacacacagcagcaaac 1546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1250 GTTGACAAAGACGGTTCCTCCAGTCCCGAGCGGAGGATTCCTACTGACTGAGAGCAAC 1309

QY 1547 cagcagagatgcgcaaa-----catctctaccag 1573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1310 AATGAGGAGGCTTCAGCGGGGTAAGCGGAACAGCAGCCACAGGTGTCTCATCTACCTG 1369

QY 1574 caaagccacgtgtctctccccccagcccgcaatggagtgactcttctgaaggagggtccct 1633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1370 ACCAACACCATGCTCCAGGTATGAGAAATGGAGGCTGCCAGGGGTGAAGGAAGAACAA 1429

QY 1634 ---cgctcttttgaactcctcaagcccccctcccatctgcctgag-----ggac 1678
```

```
Db 1430 CAGCGCATTTTGAGCGTTTGGCAGCAGCAGGAATGGATTGAGTATAGCGTCATCAGAA 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1679 tccatcaaagtgtatcaacaagaaggaggtgagtgatgtgtttcgcagtgtgaccactgc 1738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 GGATTTAAGTCTCTGAGTGGAGATGGAGAAGAACTGAGGGCGTACCGCTGTATCCACTGC 1549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1739 cactctcttctcctagattatgtattccaccatccacatggggtgcacatgtttcct 1798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1550 AGAGTTTCTTCTCGATCATGTATACACCATCCACATCGGCTGTATCGCTTCCGA 1609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1799 gatcctttgagtgtaacatgtgtgctatcgaaacccacgacgtcgtatgattctctct 1858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1610 GACCCCTTGTAGTGAACCTATGCGGTACCGAGTACCGAGCCGTTATGAGTTCATCG 1669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1859 cacatcgccagaggagagacag 1881
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1670 CACATCACACGTGGAGAGCACCG 1692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AB017615
LOCUS AB017615 2688 bp mRNA ROD 09-APR-1999
DEFINITION Mus musculus mRNA for Eos protein, complete cds.
ACCESSION AB017615
NID 94062982
VERSION AB017615.1 GI:4062982
KEYWORDS Eos protein.
SOURCE Mus musculus (strain:ICR) Newborn cDNA to mRNA.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Homma,Y., Kiyosawa,H., Mori,T., Oguri,A., Nikaido,T., Kanazawa,K.,
Tojo,M., Takeda,J., Tanno,Y., Yokoya,S., Kawabata,I., Ikeda,H. and
Wanaka,A.
TITLE Eos; a novel member of the Ikaros gene family expressed
predominantly in the developing nervous system
JOURNAL FEBS Lett. 447 (1), 76-80 (1999)
MEDLINE 99232954
REFERENCE 2 (bases 1 to 2688)
AUTHORS Wanaka,A.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-1998) to the DDBJ/EMBL/GenBank databases. Akio
Wanaka, Fukushima Medical College, Institute of Biomedical
Sciences, Department of Cell Science; Hikarigaoka 1, Fukushima,
Fukushima 960-1295, Japan (E-mail:wanaka@cc.fmu.ac.jp,
Tel:81-24-548-2111(ex.2800), Fax:81-24-549-8898)
FEATURES
Location/Qualifiers
source
1..2688
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/dev_stage="Newborn"
753..2354
/gene="Eos"
753..2354
/gene="Eos"
/function="Ikaros family predominantly expressed in the
developing nervous system"
/codon_start=1
/product="Eos protein"
/protein_id="BAA36213.1"
/db_xref="PID:d1037199"
/db_xref="PID:g4062983"
/db_xref="GI:4062983"
/translation="WESTFCSSGDSLEKSLGAPVCPVSYPNSOHSPPSRISLAN
SIKVMYDSESSRLLDKDDSVIYVEDLSLEPLGCDGSGPPHSGIRLP
NGKLVCDVGMVCIQPNVIMVHKRSHGTGERFHCNQCCAGFTQGNLRLHKLHSEK
PFKPCFNACRRDALTGLRHTSVSSPTKPKYKNCYGRKYQOSTLEHKEKCH
NYLASLSDAQALTQCPDRLMVPDMSLHPSTERTPTDIRLANSLTRKRSTPO
KFVEGKQMRFSLDLPYDVNAGGVEKDELVAHHLGPEPGGSLAFVGTBHLRLRL
PPTNCISELTPVISSVYTMQPIPSRLELPSRENGEGPEDLDGGGPLLYRAGSLTD
```

```
/product="ikaros homolog"
/protein_id="AAB53434.1"
/db_xref="PID:g2062740"
/translaton="MEMEEAQEMSOMPRDPPNDVSEENDEAMPIDELSSANLQ
HNNGDRGELRPFQCTOGKAGSTOKNLLRHITKLSGKPKFKHLNACRRRAL
SGHLRTHSVKPKHCAVCGRSYKORSLEEKHERCHNYLOCMGLONSIIYVVEESNQ
NEQEDLSQMSKRALVLDRLANNVARKSTMPQKFGVERFNSISFEGGPELMOPH
VIDQAINSAINYLAES3PLIQTSPSSDMVNGSMYPLHKPPAEGHGLSKDSAAE
NLLAKRSKSSKESKDSHSGODSTDSNEEKAGVAGSLIYTNHITSGVRNG
VLPLVKEQORQYAMRASIEIASGEKFLVSIGEEOVRAVRCHECHRLFLDHMYTTH
MGCHGRDPFECNLCGRHSODRVEFSSHMTRGHRH"
BASE COUNT 638 a 517 c 488 g 539 t 1 others
ORIGIN

Query Match 14.6%; Score 290.6; DB 4; Length 2183;
Best Local Similarity 57.2%; Pred. No. 8.2e-56;
Matches 648; Conservative 0; Mismatches 419; Indels 66; Gaps 4;

Qy 794 ggcgaacccgcgttcagatgaatcagtcgaggggcatctttactcagaagtgaaacctc 853
Db 212 GGAGAGCGTCCAGTCCAGTGCACCGAGTGGCGCTTCTTCCACCCAGAAAGCAACCTG 271

Qy 854 ctccgtcaataaactcacacggggggaacaaaccttttaagtgtcacctctgcaactac 913
Db 272 CTACGTCTACATCAAGTCCATCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331

Qy 914 gcatcgaagagagatgcgtcagcgagcaccttagcacacattctgtgagaaagcg 973
Db 332 GCCTGTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391

Qy 974 tacaagtgtgagttctcgggaagagctacagcagagaagctccctggagagcacaag 1033
Db 392 CACAAGTGTCTACTGTGGCGGAGCTACAAGACGCTAGTCTCTCGAGGAGACACAAG 451

Qy 1034 gaacgtccgcagctttcttcagaacccctgacct----- 1069
Db 452 GAGCGGTGTCAACACTCTCAGTGCATGGGCTCCAGAGACAGCATCTATACAGTAGTA 511

Qy 1069 --gggggagcgtgcaagtgtgaggaagacacacatacaagcgagatggagtgagaga 1126
Db 512 AAGGAAGAAGCAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 571

Qy 1127 gctctgctcggacagattagcaagcaatgtggttaagcgaagaaagctcgatccctcag 1186
Db 572 GCCTTGGTGTAGACAGACTAGCTAATAATGTAGCCAAACGTAAGAGACACTATGCCAC 631

Qy 1187 aaattcatcgtgagaagcgacgtcttcgatgccaactacaatcccggtacatgtac 1246
Db 632 AAGTTTGTAGGTGAGAAACG-----CTTCTCAACATCTCCTTCGAG 673

Qy 1247 gagaaggagaacgagatgatgcagaccgagatgagaccagcgaatcaataacgcaatc 1306
Db 674 GGTGCGCCGGGGAACTGATCGAGCTCATGTCATCGACCGAGCCATCAAGTGCCTATC 733

Qy 1307 agctatctaggggtgaagccttcgcgccttagtccagatccgcctgctcccaacctct 1366
Db 734 AACTACCTGGTGGCGAGTCACTCCGACCGCTCATCCAGAGCTCCCAACCTCCTCTGAC 793

Qy 1367 gagaaggtcccaagtcacagcagtgatgaccccatagcaacttactcggcgcatatgcca 1426
Db 794 ATGGGGGTGATGGGCTCCATGTACCCCTCCACAGACGCCCTTCGAGAGGGCGGCGCTG 853

Qy 1427 atgggggcccgcagagatgagaaagaaagcgaatccctcctgccagagagatcttgct 1486
Db 854 TCAGCCAGGAGACAGCGCAGCGCAAAATCTTGTGTGCTCGCAAACTCCAAGTCAGCTCC 913

Qy 1487 tctgaacagaggtctgtcccccaataacagtgccagagactcacagacagcagcaac 1546
Db 914 AGCGAAGAGGAGCGGTCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 973

Qy 1547 cacagga-----tcgcaacatctctaccagcaagcgaagcgaagtggtctc 1591
```

```
Db 974 AATGAAGAGAGAGCAGGTGCGGGGCTCAGGCTCATCTACTGACAAACCATCAGC 1033
Qy 1592 cccagggcccgcaatggatgctctcttga-----aggaggtccctcgctctttttaa 1645
Db 1034 TCCGGGTACGTAATGCGGTGCTGCCCTGGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1093
Qy 1646 ctctcaagcccccctccactctgctgagggactcccaagtgcatacaaaagaaggg 1705
Db 1094 GCCATCGGGGCCAGTATCGAGATTCCCTCGAGGGGTTCAAGGTGCTGAGTGAAGAGG 1153
Qy 1706 gaggatgagtggtttctgagtgagacactgccaactctctctctagattatgtatg 1765
Db 1154 GAGAGGTGAGGGCATATCGTGTGAACATGCGCATCTCTTCTTCCATAGACCATGTGATG 1213
Qy 1766 ttaccatcacatcggggtgcatggtttcctgtagtccctttgagtgatgagtggtg 1825
Db 1214 TACACCATCCACATGGCTGCCATGGCTTCGGAGACCCCTTCGAGTCAACCTCTGTGGC 1273
Qy 1826 tatgaagcagcagctgctgagttctctctctcactcaccatccagagagagagca 1878
Db 1274 CATGCCAGTCAGGACCGTTACGAGTTTTCGTTCACACATGACCCGAGGGAGCA 1326

RESULT 9
AF092175 2309 bp mRNA VRT 29-SEP-1998
LOCUS Danio rerio ikaros mRNA, complete cds.
ACCESSION AF092175
NID G3661582
VERSION AF092175.1 GI:3661582
KEYWORDS zebrafish.
SOURCE Danio rerio
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Rasbora; Danio.

REFERENCE 1 (bases 1 to 2309)
AUTHORS Anemiyu, C. and Kawasaki, H.
TITLE Characterization of zebrafish ikaros, a gene necessary for differentiation of the immune system
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2309)
AUTHORS Anemiyu, C. and Kawasaki, H.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Center for Human Genetics, Boston University School of Medicine, 715 Albany Street, Boston, MA 02118, USA

FEATURES
source Location/Qualifiers
1. 2309
/organism="Danio rerio"
/db_xref="taxon:7955"
86..1699
/codon_start=1
/product="ikaros"
/protein_id="AAC61763.1"
/db_xref="PID:g3661583"
/db_xref="GI:3661583"
/translaton="MEMEEAQEMSOMPRDPPNDVSEENDEAMPIDELSSANLQ
HNNGDRGELRPFQCTOGKAGSTOKNLLRHITKLSGKPKFKHLNACRRRAL
SGHLRTHSVKPKHCAVCGRSYKORSLEEKHERCHNYLOCMGLONSIIYVVEESNQ
NEQEDLSQMSKRALVLDRLANNVARKSTMPQKFGVERFNSISFEGGPELMOPH
VIDQAINSAINYLAES3PLIQTSPSSDMVNGSMYPLHKPPAEGHGLSKDSAAE
NLLAKRSKSSKESKDSHSGODSTDSNEEKAGVAGSLIYTNHITSGVRNG
VLPLVKEQORQYAMRASIEIASGEKFLVSIGEEOVRAVRCHECHRLFLDHMYTTH
MGCHGRDPFECNLCGRHSODRVEFSSHMTRGHRH"
BASE COUNT 656 a 508 c 572 g 573 t
ORIGIN
```

LHSGKPKPCFSCVACRRRDALTGHLRTHSVGKPHKCNVCGRSKQKSSLEHKERC
HNYLQVSHMAAGQVSHVPPMEDCKQEPIMDNNISLVPEFPAVLEKTFANMGR
KSTPQKFEVGEKLMRFSPDIHFDMNLTYEKEAEMLQSHMDQAINNAITVYGAALH
PLMOHAPSTIAEVAPISSAYQVYHPNRIERPIRETSDSHENMDGPTSLIRPKSR
POERASPSNCLDSTDESSHDSDSQYOGNPNKPKQSPAYMKEDVKALDATKAP
KGLADIKYVFNGEQEIIRAFKCEHCVFLDLHVYTIHMGCHGYRDPLECNICGYS
ODRYBFSSHIVRGHTFH

misc_feature

517..834

/note="encodes N-terminal zinc finger domain"

misc_feature

1594..1746

/note="encodes C-terminal zinc finger

domain/hunchback-like zinc fingers"

BASE COUNT 534 a 466 c 445 g 402 t

ORIGIN

Query Match 16.7%; Score 331.4; DB 12; Length 1847;
Best Local Similarity 58.0%; Pred. No. 5.2e-65;
Matches 733; Conservative 0; Mismatches 446; Indels 84; Gaps 5;

QY 701 agcagaccacacagtggaagatgaactgcagctgtgcgggtatcttcgcatagcttc 760
DB 490 ATCCGGCTTCCGAATGGTAAATGTGACGTCTGTGGCATGTTTGCATTGGGCC 549
QY 761 aacgtcttgatgttcataagcgaagccataccgcggaacccgcttcagtgtaacag 820
DB 550 AATGTGCTTATGTATATAAAGAGTACACCTGTGGAGCGCCCTTCACCTGAAC 609
QY 821 tgcggggcatctttactcagaaggttaacctctcctcgtctaatattaaactgcacacgggg 880
DB 610 TCGGAGCTTCTTTTACCACAGAGGCAACCTTCTGAGACACATAAAGTTACACTCTGGA 669
QY 881 gaaaac 940
DB 670 GAGAACCCCTTCAATGTCTTTCTGTAGCTATGCTGTAGAGAGAGGACGCTCTACA 729
QY 941 ggacacttagacacatctgtgagagcgcgtacagtgtagtgagttctgcggaagac 1000
DB 730 GGACACTCAGGACCCATCTGTGGGTAAACCTCAAGGTAACTACTGTGGCCGAAGC 789
QY 1001 tacaagcagagaagctccctggaggagcacaaggaacgctgcgagctttcttcagac 1060
DB 790 TACAAGCAGCGCAGCTCACTGGAGGAACACACAGGACGCTGTACACACTATCTCCAGAT 849
QY 1061 cctgacctgggggaagc-----tgcaagctgtg 1088
DB 850 GTCAGCATGGAGCTCCGGGAGGTCATGAGTCACCATGTACCGCTATGGAAGATTGT 909
QY 1089 aggcaga-----cacatcaagcgcagatgggaagtggaagagctctc 1132
DB 910 AAGAACACAGAGCTATCATGGACACAAATATTTCTGTGTGCTTTTGAGAGACCTGCT 969
QY 1133 gtctggacagattagcaagcaatgtggctaaagcgaagcctcgatcgctcgaattc 1192
DB 970 GTCATAGAGAAGCTCAGCGCAATATATGGGAAGCGCAAAAGCTCCACTCCTCAGAAATT 1029
QY 1193 atcgggtagaagcgcactgcttc-----gatgcgaactacacatccccggtacatg 1243
DB 1030 GTGGGGGAAAAGCTTATGCGATTACGCTACCCAGATATTCATTTGTATGAACTTAAAC 1089
QY 1244 tacgagaagagacagatgatgcagaccggatgatgacacacacacacacacacacac 1303
DB 1090 TATGAGAAGAGGCTGAGCTGATGAGTCTCATATGATGACCAAGCCATCAACAATGCA 1149
QY 1304 atcagctatagggctgaagccttcgcgccttagtcagactcgcctgctccacc 1363
DB 1150 ATCACTACCTTGGAGCTGAGGCCCTTCACCTCTGTATGAGCATGCACCAACCAATC 1209
QY 1364 tctgagatggtccacgcatcagcagtggttaccacacacacacacacacacacacacac 1405
DB 1210 GCTGAGGTGGCCCGAGTTATAAGCTCAGCTTATTTCTCAGGTCTATCATCCAAACAGGATA 1269
QY 1405 -----acttactcggccgatatgcceatggggcccccaggagatggaagaaacacgg 1459

DB 1270 GAAAGACCCATTAGCAGGGAACATCTGATGTCACGAAAACATGATGGCCCCACT 1329
QY 1460 atccctctgcagagaagatcttctctgaacagtgctgtccccaataaacagtcgc 1519
DB 1330 TCTCTCATCAGACCAAGAGTCGACCCAGGAAAGAGAGGCTTCGCCAGCAATAGCTGC 1389
QY 1520 caggactccacagacacccagcagcaaccacacagagatcgcccaacatctctaccagcaagc 1579
DB 1390 CTCGATTCTACTGACTGAGAAAGTAGCCATGATGACCGC---CAGTCTCTACCAAGGAAC 1446
QY 1580 cagctggctctcccccagccgcgaatggatgctctcttctgaagagagtcctcgcctc 1639
DB 1447 CTGCTTAAATCCACAGAGGAACAAGCCAGCTTACATGAAGGAGATGTCGAAGCT 1506
QY 1640 ttgaaactctcaagccccctcccatctgcctgagggactccatcaaaagtcatcaacaaa 1699
DB 1507 TTGGATGCTTACCAAGGCCCCCAAGGCTCTCTGAAGGACATCTATAAGGTTTCAATGGA 1566
QY 1700 gaaggggagtgatgagtggtttcgtatgtgaccactgccacgctcctctctcattat 1759
DB 1567 GAAGGAGAACACATGAAGGGCTTCAAGTGTGAGCACTGCCGAGTCTCTTTCTAGACCAT 1626
QY 1760 gtgagtccacatccatcaggtgggtgcccaggtttccctgacatcccttgagtgtaacatg 1819
DB 1627 GTCATGTACACATTTCATCATGGTGGCTACCGGACCCACTGGAATGCAACATC 1686
QY 1820 tdtgctatcgaagcagcagatcgctatgattctctctcctcgcacatgccagagagagcac 1879
DB 1687 TGTGCTCAGAGAAGCCAGGACCGCTACGAAATTTTCATCACACATTTTCGAGGGGAGCAC 1746
QY 1880 aga 1882
DB 1747 ACA 1749

RESULT 8

OMU92200

LOCUS

OMU92200 2183 bp mRNA VRT 02-MAY-1997

DEFINITION

Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, Ik-2 isoform, complete cds.

ACCESSION

U92200

NID

G2062739

VERSION

U92200.1 GI:2062739

KEYWORDS

rainbow trout.

SOURCE

Oncorhynchus mykiss

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

1..2183

/organism="Oncorhynchus mykiss"

/strain="shasta"

/db_xref="taxon:8022"

/cell_type="thymocyte"

50..1336

/gene="Ikaros"

50..1336

/note="Ik-2 isoform; similar to mouse and human

Ikaros/lyf-1; alternatively spliced form missing exon 3"

/codon_start=1

gene 1. 2496
/gene="Ikaros"
CDS 81.1649
/gene="Ikaros"
/note="Ik-1 isoform; similar to mouse and human Ikaros/LyF-1; Ikaros/LyF-1 homolog"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB97474.1"
/db_xref="PID:g2062742"
/db_xref="GI:2062742"
/translation="MEMEAEQMSQPRDPPNDVSEENDEAMPIDLSASNLQ HNRGDEGLACNKKVEARCDENGLAIDMMNGEEREECAEDLRVLDASQKVNQSH AGPDSKGPSSAGGIRLPNGKIKDCIGVICVGPVLYVHKRSHGRFQCTQCGA SFTQGNLLRHKLHSGEKFCHLCNRYARRRDAUSGHLRHSVGRPHKCAVCGRSY KQRSLSEERHCNRYLQCMGLQNSIYTVVEESQNEQREDLSQMSKRALYDLRLA NNVAKRSTMPQKPFVGRFNTSFEFGGPELMOPHYIDQAINSAINLYGAESLRPL QTSPTDSMGVMGSMYPLHKPAPAEHGLSKADSAENLLAKSKASSSEKDGSPHS QDSTPTDSNNEKAGVGASGLLYLTHNITSGVRNGVLPVKEQOROXYEAMRASLEI ASEGFVLSEGEQVRAIRCEHCRILFLDHWMTIHMCHGFRDPFECNLCGRHSQDR YEFSSHMIRGEHRT"
BASE COUNT 707 a 592 c 597 g 599 t 1 others
ORIGIN

Query Match 16.9%; Score 335.6; DB 4; Length 2496;
Best Local Similarity 58.0%; Pred. No. 66;
Matches 711; Conservative 0; Mismatches 449; Indels 66; Gaps 4;

Qy 701 agcagaccaacagcagtggaagatgaactgcagcgtgtgcggttatctcgtcgttagcttc 760
Db 432 ATCCGCTTCCACCGGGAAGCTCAAGTGTGTATCTGTGGGATAGTTTGGATTGGCCCC 491
Qy 761 aacgtcttgatggttcataaagcagaccatcccgcggaacgcccgttcacgtgtaatacag 820
Db 492 AATGTGCTGATGTCACAAAGGAGCCACACTGGAGAGCGCTCCATTCCAGTCACCCAG 551
Qy 821 tgcggggcatcttttaactcagaagaagtaaacctctcctcctcctcattataaactcacacagggg 880
Db 552 TGTGGCGCTTTCTTACCAGAAAGGCAACCTGTCTACGTACATCAAGCTCCACTCAGGA 611
Qy 881 gaaaaacctttaagtgtcacctctgcactacgcatacgcatacgaagagagatgcgctcacg 940
Db 612 GAGAAGCCCTTCAAGTGTACCTGTGCAACTATGCTCTGCCGGAGAGAGCCCTCAGC 671
Qy 941 ggacaccttagcacacattctgtgagagccgtacacagtgatgtgctgcgggaagc 1000
Db 672 GGTCACTGCGTACCCACTCTGTTGGAACCCCAACAGTGTGCTTACTGTGGGGAGC 731
Qy 1001 tacaagcagagaagctccctcgagagagacagaagacgctcgcagctttcttcagaac 1060
Db 732 TACAAGCAGCTAGTCTCTGAGGAACACAAAGAGCGGTGTACAACTACTCTCAGTGC 791
Qy 1061 cctgacct- 1093
Db 792 ATGGGGCTCCAGACAGCATCTATACACTAGTAAGGAAGAAAGCAACCAAGATGAGCAG 851
Qy 1094 agacacatcaaacgcagatgggaagtgcagagagctctcctcgctggagacattagcaagc 1153
Db 852 AGGGAAGACTTAAGCCACATGGATCTAAGAGAGCGCTTGGTGTCTAGACAGACTAGCTAAT 911
Qy 1154 aatgtggttaagcgaagaaagctcgtgcctcagaaattcatcgtgtgagaagcgcactgc 1213
Db 912 AATGTAGCAACGTAAGACGACTATGCCACAGAGATTTGTAGGTGAGAAACGCTTCTCC 971
Qy 1214 ttcgatgccaaactcaatcccggtctacatgtacgagaaggaagacagatcatcagacc 1273
Db 972 AACATCTCTTTCAGGGTGGCCCC 1013
Qy 1274 cggatgatgaccagcagcatcaataaacgccctcagctatctataggggctgaagccttcgcg 1333
Db 1014 CATGTGTCGACGAGGCCATCAACAGTCCCATCACTACCTGGTGGTGGAGTCACTCCGA 1073

Qy 1334 cccttagtcagactcgcgctgtctccacactctgagatgggtccagtcacagcagtg 1393
Db 1074 CCGCTCATCCAGACCTCCCAACCTCCTGTACATGGGGGTCTATGGGCTCCATGTACCCC 1133
Qy 1394 taccatagacttactcggcgcatatgccaatggggggcccgagagagagagaaag 1453
Db 1134 CTCCACAAGCCCTCGAGAGGCCAGCGCTGTACGCCAAGGACAGCGAGCGAAT 1193
Qy 1454 aaacggatctctcgtccagagaagatcttgcctcttgaacgaggtgtgtcccccaatac 1513
Db 1194 CTTCTGTGTCGCAAGTCCAAGTCAGCTCCAGCGAGAAGAGCGGCTCCCCCAGCCAC 1253
Qy 1514 agtgcagagactccacagacacagcagacagcaaacacagagatgcgcaaatctctaccag 1573
Db 1254 AGTGCCAGGATTCCACCGACAGGAGCAACAATGAAGAGAAGGAGGAGGTGTGCGGGGCC 1313
Qy 1574 caaagcacactggttcctccccc-----ggcccgcgaatggatgctctt 1618
Db 1314 TCAGGCTCTACTCTACCTGACAAACACATCAGCTCCGGGTACGTAATGGCGTCTGCC 1373
Qy 1619 ctga-----aggaggtccctcgtcttgaactcctcaagccccctccatctgctg 1672
Db 1374 CTGTGAAGGAGGAGCAGCAGCGCAGTACGAGGCCATCGGGCCAGTATCGAGATTGCC 1433
Qy 1673 agggactccatcaagtgcatacaaaagaagggaggtgagatggtgttcgatgtgac 1732
Db 1434 TCGAGGGGTTCAGGTGCTGAGTGGAGGGGGAGCAGGTGAGGCATATTCGCTGTGAA 1493
Qy 1733 cactgcaacgctctcctccatgattatgattcaccatccacatgggggtggtgcaaggt 1792
Db 1494 CACTGCGGCATCTCTCTCCATGACCATGTATGTACACCATCCACATGGCTGCCATGGC 1553
Qy 1793 ttccgtgaccccttgagtgtaacatgtgtggtctatcagaagccacgacgtatgattc 1852
Db 1554 TTCGAGACCCCTTCGAGTGCACCTCTGTGGCATCGCAGTCAGGACCGCTTACGAGTTT 1613
Qy 1853 tctctcaacatcgccagagagagaca 1878
Db 1614 TCGTCACACATGACCCGAGGGGAGCA 1639

RESULT 7
AF044257 1847 bp mRNA ROD 04-FEB-1998
LOCUS Mus musculus multi-zinc finger protein helios mRNA, complete cds.
DEFINITION AF044257
ACCESSION AF044257
NID 92829276
VERSION AF044257.1 GI:2829276
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1847)
AUTHORS Hamm,K., Cobb,B.S., McCarty,A.S., Brown,K.E., Klug,C.A., Lee,R., Akashi,K., Weissmann,I.L., Fisher,A.G. and Smale,S.T.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) HHMI/McI, UCLA, 675 Circle Drive South, Box 951662, Los Angeles, CA 90095-1662, USA
FEATURES
source 1.1847
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS 178.1758
/note="Ikaros family member"
/codon_start=1
/product="multi-zinc finger protein helios"
/protein_id="AAC00513.1"
/db_xref="PID:g2829277"
/db_xref="GI:2829277"
/translation="METDAIDGYITCDNELSPGEHANMAIDLTSTPNGQHASPSHM TSPNSVKLEMSDECDROPLEIRGHDESSLEPIESSEVADNRKVDLOGE GTIRLPNGKLKCDVCGMVGIPNLVHVRSHRTPFPFCNCGGASFTQGNLLRHK

QY	1370	atggtccacgtcatcagcagtggtgtaccccatagcacttactcggggccgatgccaatg	1429
Db	786	GTGGTGCCAGTCATCAGCTCCATGTACCAGCTGCACAAGCCCCCTCAGATGGCCCCCA	845
QY	1430	ggg-----gccccgagagatggaagaacacgactcctcctgccagagaag	1477
Db	846	CGGTCAACCATTCAGCAGGACCGCTGGATACTTGTCTGCTGTCTCCAAAGGCCAAG	905
QY	1478	atcttgcttctgaacaggtctgtcccccaataaacagtgcccagagactccacagacc	1537
Db	906	TCTGTGTCTATCGGAGGAGGCTTCCCGAGCAACAGCTGCCAAGACTCCACAGATACA	965
QY	1538	gacagcaaccacgagg---atcgccaacatctctaccagcaagaagcgaagtgtctctccc	1594
Db	966	GAGAGCAACCGCGGAGGAACAGCGAGCGGCTTATCTACTTAACCAACACATCAACCGG	1025
QY	1595	caggcccgcaatggagctcttctgaagaggtccctcgtctcttcttgaactctcctcaag	1654
Db	1026	CATGCAACGCAATGGCGTGGCTCTC---AAGGAGGAGCAGCGCGCTTACGAGGTGCTGAGG	1082
QY	1655	ccccctcccatctgctgagggactccatcaaaagtgtatcaacaaagaagggaggtgatg	1714
Db	1083	CGGGCTCAGAGAACTCGCAGGATGCTTCGCTGTGGTCAGCAGCAGTGGCGAGCAGCTG	1142
QY	1715	gatgtgttcgatgtgacacactgcacgtctctctctctctctctctctctctcaccatc	1774
Db	1143	AAGGTGTACAAGTGGCAACACTGCGCGGTGCTCTCTCTGGATCAGCTCATGTATACCAT	1202
QY	1775	cacatg-----gggtgccatggttccctgacatcccttggatgaactgtgtgac	1825
Db	1203	CACATGGGTGCATGGCTGCCATGGCTTTCGGGATCCCTTTGAGTGTAACTATGTGTG	1262
QY	1826	tatcgaagcgaacgacgtcgtatgattctctctcactcgcagagagagacagagcc	1885
Db	1263	TATCAGCCAGCAGCAGGTACGAGTTCTCATCCCATATCACGCGGGGAGCATCGTTAC	1322
QY	1886	atgttgaagtggc	1899
Db	1323	CACCTGAGCTAAAC	1336
RESULT	6		
OMU92201			
LOCUS	OMU92201	2496 bp	mRNA
DEFINITION	Oncorhynchus mykiss Ikaro homolog (Ikaro) mRNA, Ik-1 isoform, complete cds.		
ACCESSION	U92201		
NID	92062741		
VERSION	U92201.1	GI:2062741	
KEYWORDS	rainbow trout.		
SOURCE	Oncorhynchus mykiss		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.		
REFERENCE	Hansen, J.D., Strassburger, P. and Du Pasquier, L.		
AUTHORS	Conservation of a master hematopoietic switch gene during vertebrate evolution: isolation and characterization of Ikaro from teleost and amphibian species		
TITLE	Eur. J. Immunol. 27 (11), 3049-3058 (1997)		
JOURNAL	98056818		
MEDLINE	2 (bases 1 to 2496)		
REFERENCE	Hansen, J.D.		
AUTHORS	Direct Submission		
TITLE	Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland		
JOURNAL	1..2496		
FEATURES	/organism="Oncorhynchus mykiss" /strain="Shasta" /db_xref="taxon:8022" /cell_type="thymocyte"		
source			

MEDLINE	93068267	Location/Qualifiers
FEATURES	1..1550	/organism="Mus musculus"
source	/db_xref="taxon:10090"	/cell_type="T cell"
	/dev_stage="adult"	/tissue_lib="Lambda ZAP EL4"
gene	39..1334	/gene="Ikaro"
CDS	39..1334	/gene="Ikaro"
	/codon_start=1	/product="Ikaro DNA binding protein"
	/protein_id="AAA66193.1"	/db_xref="PID:g198287"
	/db_xref="GI:198287"	/translation="MDVDEGQDMVSQVSGKESPPVSDTPDEDEPMPVPEDLSLTSQAQ
		QNSKSDRGMEPFOCNOCGASFTQKGNLRLHKLHSEKPKFKCHLNCYACRRRLALT
		GHLRHSVKGPKHGKCYGRSYKORSLEHKEKRNHLESMLPGVPVPIKEETNHNE
		MAEDLCKGAERSLVLDRLASNVAKRKSMPQKFLGDKLSDMPYDSANYEKEDNMTS
		HYMDQAINNAIYLGAESLRPLVOTPPGSSEYVPISSMYOLHKPPSPGPPRSHSAQ
		DAVDNLLLSKAKSYSSEREASNSCODSTDESNAEERSLGLIYLHINPHARNG
		LALKEORAYEVLRAASENSODAFRVSTSGEQLKYKCEHCRVLFDLHVMTYIHGC
		HCCHFRDPFECNMCGYHSQDRYKESHITRGEHYHLS"
BASE COUNT	385 a	431 c 423 g 311 t
ORIGIN		
Query Match	17.5%;	Score 348; DB-12; Length 1550;
Best Local Similarity	62.0%;	Pred. No. 9.5e-69;
Matches 716; Conservative	0;	Mismatches 375; Indels 63; Gaps 8;
QY	794	ggcgaacgcccttcacgttaacatcagtcggggcattctttactcagaaagtaaacctc 853
Db	198	GGTGAACGGCCTTTCAGTGCAACCACTGTTGGGCTCTCTTTACCCAGAAAGCAACCTC 257
QY	854	ctccgtcataataaactgcacacgggggaaaaaaccttttaagtgtcacctctgcaactac 913
Db	258	CTCGGCACATCAAGTGTCTACTCGGTGAGAAGCCCTTCAAAATGCCATCTTTGCAACTAT 317
QY	914	gcatacgaagagagatgcgtcagggacacttagcacacatctctggagagacgcg 973
Db	318	GCTGCGCGGAGGAGCGCCTCACCGGCACCTGAGGACGACCTCCGTGTGTAAGCCT 377
QY	974	tacaagtgtgattctgcgggaagaatacagaacagagaagctccctggaggagcacaag 1033
Db	378	CACAAATGTGATATTGTGCCGGAGCTATAAAGACGGAAGCTCTTTAGAGGAGCATAAA 437
QY	1034	gaacgtgcgagcttttcttcagaaccctgacctgggggagc-----ctgcaagtgtg 1087
Db	438	GAGCGATGCCACACTACTTGGAAAGCATGGGCTTCCGGCGGTGTGCCCGATTAAG 497
QY	1088	gaggcaagacacataaag-----ccgagatgggaagtgcagagact 1129
Db	498	GAAGAACTAACCCACACGAGATGGCAGAGACCTGTGCAAGATAGGACGAGAGAGTCC 557
QY	1130	ctcgtctgcacagatagcaagaatgttgtaagcgaaaaagctcgtatgcctcagaaa 1189
Db	558	CTTGTCTCTGACAGGCTGGCAAGCAATGTGCCAAACGTAAGAGCTCTATGCTCCAGAAA 617
QY	1190	ttcatcgggtgagaagcgccactcttcgatgccaactacaatcccggtctacatgacag 1249
Db	618	TTTCTTGGAGACAAGTGCCTGTGACAGATGCCCTATGACAGTGCCAATATGAGAAGGAG 677
QY	1250	aaggagaacagatgatgcagacccggatgatggaccacacatcaataaagccatcagc 1309
Db	678	-----GATATGATGACATCCCACGCTGATGGACGAGCCCATCAAAATGCCATCAAC 728
QY	1310	tatctaggggctgaagccttccccccttagctagtcacactccgctctccacactctgag 1369
Db	729	TACCTGGGGCTGAGTCCCTGCGCCCATTTGGTCAGACACACCCCGCT---AGCTCCGAG 785

TITLE Direct Submission
JOURNAL Submitted (12-MAR-1997) J.P. Liippo, Turku Immunology Centre and Department of Medical Microbiology, Turku University, Kiinamyllynkatu 13, FIN-20520 Turku, FINLAND

FEATURES
source
1. 1557
/organism="Gallus gallus"
/strain="RPL1"
/db_xref="taxon:9031"
/germline
/tissue_type="thymus"
/cell_type="T lymphocyte"
/clone_lib="PCDM8 thymus cdna"
1. 1557
/gene="Ikaros"
1. 1557
/gene="Ikaros"
/codon_start=1
/product="Ikaros transcription factor"
/protein_id="CAA72531.1"
/db_xref="PID:e309211"
/db_xref="PID:g2330595"
/db_xref="GI:2330595"
/db_xref="SWISS-PROT:O42410"
/translation="METDRAQDMQSVGKESPPISDVPDDADPMVPVDELSTTTGCG QSVNERYLAGNIKIETQSDENGRACMECEADRLMDLSDKXMGSHNGPGSK AMSGVGRLPENGKLDICGICIGPNVLMVHNRSHGTGERPQCNGGASFTOKGNL LRHLKHSRPFKHLCLNACRRDLTHLRHSVGHKPKGCGYKQKRSLEE HKERCHNYLQKLSISNLYSVIKEETNOSEMAEDLKIGERSLVLDORLASNVAKRS SMPKQFVEKCLSDLPYDNYEKENIMOTHVIDOAINNAISYLGAESLRPLVOTP PVGSEVPVISPVMYQLHKLPGDNTSRNHTAQDSAVENLLLSRAKSVSBERDASPEN SCQSDTFESNNEKGLIYLTNIGHGPHARGNISVKEESROFDVLRGATONSQDAFKV ISSNGEQVRYNKEHCRVFLFDHVMYTIHMGCHGFRDPFECNMGYHSQDRYEFSSH TRGHRHMS"

BASE COUNT 456 a 368 c 387 g 346 t
UNIGIN

Query Match 20.1%; Score 399.6; DB 4; Length 1557;
Best Local Similarity 62.0%; Pred. NO. 2.3e-80;
Matches 759; Conservative 0; Mismatches 414; Indels 51; Gaps 6;

QY 700 cagcagaccacacagtggaagatgaactgcagctgtgctggttattctgcttagctt 759
DB 327 CATTCGACTTCTAATGGAAGCTAAATGTGATCTGTGGGTAATTTGGCATCGGTC 386

QY 760 caacgtcttgatggttcaaaagcagccataccgagcagacccggttccagtgtaatca 819
DB 387 CAATGTGCTCATGGTTCACATCGAAGCCATCTCGAGAACGGCCCTTCCAGTGAATCA 446

QY 820 gtgcggggcatcttttactcagaagggtaacctctcctcgtcattataaactgcacacgg 879
DB 447 GTGCGGAGCCTCTTTACGACGAAGGAAACCTCTCGGCCACATCAAGTTGCACCTCGG 506

QY 880 ggaataaccttttagtgcaccttgcactacacatcagcatgcgaagagagatgcgtcac 939
DB 507 TGAAGAAGCCCTTCAATGGACCTCTGTAACTACGCTGCGCGGACGGGATGCCCTCAC 566

QY 940 gggacaccttagcacaccttctgtggaagagccgtacaagtgtgagttctgcgggaaga 999
DB 567 AGCCACCTGCGGACTCACTCGTGTGGCAACCCCATAGTGTGATCTGTGTGCGCAG 626

QY 1000 ctacaagcagaagaacctcctggaggagcacaagaagcctccgagcttcttcaga- 1059
DB 627 CTATAAGCAGCGCAGCTCTTTGGAAGAACATAAAGAACCGCTGTATACCTACCTGCAAAC 686

QY 1059 -accctgacctggggaagcgtcaagtggtggaggaagacacacatacaagc----- 1108
DB 687 CATGAGTATCTCAAGCAATCTTTATTCAGTCATAAAGAGGAAACATAACGAGTGAAT 746

QY 1108 -----cgaatgggaagtgaagagctctcctcctggacagattagcaag 1152
DB 747 GCCTGAAGACCTGTGCAAGATAGGTCAGAAAGATCCCTCGTGTGATAGACTAGCAAG 806

QY 1153 caatgtggctaaagcgaaaaagctcgtgcctcagaataattcattcgttgagaagcgccactg 1212
DB 807 TAACGTCGCCAAACGCTAAGAGCTCTATGCTCTCAGAAATTTTGTGTGAGAAGTGTCTATC 866

QY 1213 cttcagatgcacaatacattcccggtcattgtagagaaaggaacagatgtagcagac 1272
DB 867 T---GATCTTCCATATGATGCCACCACCACTATGAGAAGGAGAACGAGATAATGCGAC 923

QY 1273 ccggtatgtagcacaagccatcaataacgcctcagctctattaggggtgaagccttcgc 1332
DB 924 CCATGTCTATACACAGCCATTACATGCCATCAGTTACCTGGGGGCGAGAGTCTTTCG 983

QY 1333 ccccttagtccagactccgctcgtcccaacctgagatggtcccaagtcacagcagtg 1392
DB 984 TCCCTCGTCCAGACTCCACCGGTT--GGTCTGAAGTGTGCTCCGCTCATCAGCCCCAT 1040

QY 1393 gtaccccatgacacttactcggcgcatatgccaattggggggcccgacaggaga----- 1446
DB 1041 GTATCAGCTCCACAAACCTCTCGGGGACATCAGACTCGCTCCCAACCATACGGCTCAGGA 1100

QY 1446 -----tggaagaagaaacggtactcctcctccagagaagatcttgcctctgaacag 1497
DB 1101 CAGCGCAGTGGAAACCTGTGTGCTCTCCAAAGCTAAATCAGTCTCTGTAACGGGA 1160

QY 1498 tctgtcccccataacagtgccccaggaactccacagacacagacacacacagagatcg 1557
DB 1161 TGCCTCCCCAGCAACAGCTGCCAAGACTCAACAGATACGGAGAGCAATAATAGGAACG 1220

QY 1558 ccaacatctctacagcaaaacagcagtggtctcctcccaagcccgcaatggatgcctct 1617
DB 1221 CAGTGGTTTAAATTACTTAAACAAACCATAGTCCACATGCAAGAAATGGCATATC--- 1278

QY 1618 tctgaagagagtgctcctcgtcttttgaaactcctcaagccctccctcctcctcctcgagg 1677
DB 1278 TGTAAGAAGAAAGCAGGCGAGTTTGTATGATTGAGGCGAGTACTGACATTTCCAGGA 1337

QY 1678 ctccatcaaatgatcaacaaagggaggtgagtgatggtgttctgagatgagccactg 1737
DB 1338 TGCTTTAAAGTGTATGACGAGCAATGGGAGCAAGTGTGAGAGTTTACAAAGTGCAGACATG 1397

QY 1738 ccagctcctctctcagatgattgattcaccatccacatgggggtgacctggttccg 1797
DB 1398 TCGAGTCTTTTGTGATCATGTGATGACAAATCCACATGGGCTGCCATGGTTCG 1457

QY 1798 tgatccctttgagtgtaacatgtggtctatcgaaagccacgctgctatgagttctcctc 1857
DB 1458 CGACCCTTTGAATGCAACATGTGTGGCTACCACAGCCAGGACAGGATGATGAATTTCTTC 1517

QY 1858 tcacatgcgcagagagagacag 1881
DB 1518 CCACATAACTCGAGGGGAGCACCG 1541

RESULT 5
MUSIKAROS
LOCUS Mouse Ikaros DNA binding protein (Ikaros) mRNA ROD 11-MAY-1995
DEFINITION L03547
ACCESSION g198286
NID L03547.1 GI:198286
VERSION
KEYWORDS DNA-binding transcription factor; Ikaros; Ikaros DNA binding protein; transcription; zinc finger protein; zinc-finger transcription factor.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1550)
AUTHORS Georgopoulos, K., Moore, D.D. and Derfler, B.
TITLE Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment
JOURNAL Science 258 (5083), 808-812 (1992)

TITLE The Ikaros gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse

JOURNAL J. Immunol. 156 (2), 585-592 (1996)

MEDLINE 96132984

GenBank staff at the National Library of Medicine created this entry [NCBI gibbsg 175401] from the original journal article. This sequence comes from Fig. 1B and C.

```

map location: /p13-p11.1.
FEATURES
  source      Location/Qualifiers
              1. .1788

```

```

source
1. 1700
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```
gene
238. .1788
/gene="IKAROS"
```

gene= INK45
/note="hik1"
CDS 238..1788

```

230...:1788
/genre="IKAROS"
/note="lymphocyte-restricted zinc finger DNA binding
protein; This sequence comes from Fig. 1C"
/codon_start=1
/product="hiki"
/protein_id="AAB50683.1"
/db_xref="PID:g1911483"
/db_xref="GI:1911483"
/translation="MDADEQDMSEFSGKSEPPVSDTPDEGDEPMIPEDLSTSGGQ
QSSKDRVSNVKTQSDENGRACMEGECAEDRLMDASGGMSHSDQSS
ALSGVGIRPNGLKCLDCIGIICGNVLMLKRRSHTEGPFQCCGASTQKGNL
LRHLKHSRGPKFLCNLCACRRARDLGLTRLSVGHKPGCGYCGYKQRTSLLE
HKERCHNYLESMGLPGTLYPIVKEETKHSEADLCKIGSERSLDRLASLNARRKS
SMPOKFLDGKSLDGLTQYIKENEMKMSHVMDQKAINNALYCAEISLPLVOTPP
GSDVTPVPMYOLHRRSEGTNRSHNSAODSAVEYLLLSLKAKLVPSESPASPSNC
QDSTVTSNNEQEQSGILYLTNIHRAAQVSLKEEHRAVDLLRAASENSQDALRVVS
TSGEOMKXVYCEHCRCRVLFLEOHVMTIHMGCGRFDRPPECNMCWGYHSQDRYEVESHHT
GERHEHNS."

```

COUNT AND			
BASE COUNT	450 a	511 c	497 g
ORIGIN			
			330 t

Query Match	Score	DB	Length
20.88	411.8	10	1788

Query Match 20.8%; SCORE 411.8; DB 10; Length 1788;
Best Local Similarity 63.38; Pred. No. 4.1e-83;

Seq. local similarity	Seq. NO.	Indels	Gaps
93.3%	4.1E-85		
Matches 777; Conservative	0; Mismatches 387;	Indels 63;	Gaps 7;

Qy 700 cagcagaccaacagtgaggaaqatqaactqcacqgtgtqcqqqtatcctqcattagctt 759

— — — — —

DB 364 CATTCGACTTCCCTAACGGAAAACCTAAAGTGTGATATCTGTGGGATCATTTGCATCGGGCC 623

Qy 760 caacgtcttgatggttcataagcgaagccataccggcgaaagccggttccagtgtaatca 819

624 603

524 CATTGCTCATGGTTCACAAAGAAAGCCACACATGGAGACGGCCCCCTCCAGTGGCAACA 883

Qy 820 gtgcggggcatcttttactcagaaaggtaacctcctccgtcataattaactgcacacggg 879

Db 684 GTGGGGGGCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCGG 743

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qy 880 ggaaaaaaccttttaagtgtcacctctgcaactacgcatgccaaaggagagatgcgctcac 939

Db 744 GGAGAAGCCCTTCAAAATGCCACCTCTGCAACTACGCCCTGCCCGCGGAGGACGCCCTCAC 803

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

999
940 gggacaccttaggacacattctgtggagaagccgtacaagtgtgagttctgcggaagaag

Db 804 TGGCCACCTGAGGACGCACCTCCGTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAAG 863

[illegible][illegible]

TITLE Cloning and sequencing of hik-1, a cDNA encoding a human homologue of mouse Ikaros/Lyf-1

JOURNAL Immunol. Lett. 49 (1-2), 139-141 (1996)

MEDLINE 96252222

REFERENCE 2 (bases 1 to 3629)

AUTHORS Niefeld, W.

TITLE Direct Submission

JOURNAL Submitted (10-NOV-1995) Wilfried Niefeld, Department of Virology, University of Freiburg, Institute for Medical Microbiology and Hygiene, Hermann-Herder-Strasse 11, Freiburg 79104, Germany

FEATURES

source Location/Qualifiers

1..3629

organism="Homo sapiens"

db_xref="taxon:9606"

tissue_type="bone marrow"

169..1728

gene="hik-1"

169..1728

gene="hik-1"

note="similar to mouse Lyf-1, encoded by GenBank Accession Number 574708; similar to mouse Ikaros DNA-binding protein, Swiss-Prot Accession Number Q03267"

/codon_start=1

/product="Ikaros/Lyf-1 homolog"

/protein_id="AAC50459.1"

/db_xref="PID:91289371"

/translation="MDADEGQDMQSVSKESPVSVDTPDEGDEPMPIDLSITSSGQ QSSKSDRVASNVKVEVTSDEENGRACEMNCEAEIDLRLDASGKNGSHRQGS LALSGVGGIRLPNGLKDCIGICIGPNVLMVHKRSHTERPFCNOCNGASFTOKGNL LRHLKHSSEKPFKHLNACRBDALTLRLHSHVSKPKHCKGCSRSYKQRSLEE HKERCHNVLESMGLPTLYPVIKETNHEMAEDLCKIGERSLVLRASNAKRKS SMPQKFLGDLGSLDTPYDSSASYEKENEMKSHYMDQAINNAILYGAESLRPLVQTP PGSEVVPVSPMOLHKLPLAETPRSNHSAQDAVENLLSLKALVPSEASPSN SCQSDTFSENEQSRGLIYLTNHIAPHARNGLSLKEHRAVDLLKLAASNSODALR VVSTSGEOMKYVKECHRCVFLDHVMYTIHMGCHFRDPPECNMGVHSQDRYEFSS ITRGEHRFMS"

BASE COUNT 917 a 1002 c 936 g 773 t 1 others

ORIGIN

Query Match 21.6%; Score 428.2; DB 11; Length 3629;

Best Local Similarity 64.0%; Pred No. 8.6e-87;

Matches 785; Conservative 0; Mismatches 388; Indels 54; Gaps 7;

Qy 700 cagcagaccacacagtggaagatgaactgcgacgtgtgcgggttatccctgattagctt 759

Db 495 CATTCGACTTCTAACGGGAAACTAAAGTGTGATATCTGTGGGATCATTTGGCATCGGGCC 554

Qy 760 caactcttgatggttatacgaagccataccgagcaccgcccgtttccagttaataca 819

Db 555 CAATGTGCTCATGGTTTCAAAAAGAACCCACACTGTGAGAACCGGCCCTTCCAGTGAATCA 614

Qy 820 gtgcggggcatcttttactcagaaggtaacctctcgtctatattaaactcacacggg 879

Db 615 GTGCGGGGCTCATTACCCAGGAAGGCAACCTGTCTGGGCACATCAAGCTGCATTCGGG 674

Qy 880 ggaataaaccttttaagtgtcaactctgcaactacgcatgcatacgaagagagatgcgtcac 939

Db 675 GGAGAAGCCCTTCAATGCCACCTCTGCAACTACGCCCTGCCCGGAGGAGCCCTCAC 734

Qy 940 gggacaccttagacacattctgtggagagccgttacaagtgtgagttcttcggaagaag 999

Db 735 TGGCCACCTGAGGAGCGCACTCGCTGGTAAACCTCACAATGTGGATATTGTGGCCGAAG 794

Qy 1000 ctacaagcagagaactccctcgtgagagcacaagaagcgtccgagcttttttcagaa 1059

ph 795 CTATAACAGCAGAGCTCTTTAGAGAACATAAAGAGCGCTGCCACAACACTACTTGGAAAG 854

Qy 1060 ccttgacct--ggggagcgtcgaagtgtggaggaagacacatcaaacg----- 1108

Db 855 CATGGGCGCTTCGGGCACACTGTACCACCTCATTAAGAAGAAACTAATCACAGTGAAT 914

Qy 1108 -----cgagatgggaagtgaagagagctctcgtctcgtgagacattagcaag 1152

Db 915 GGCAGAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGTGGACACTAGCAAG 974

Qy 1153 caatgtggctaaagcgaataagctcgtatgcctcagaaattcatcgttgagaaagcgacctg 1212

Db 975 TAAGCTGCCAAACGTAGAGCTCTATGCCCTCAGAAATTTCTTGGGGACAGGG---CCT 1031

Qy 1213 ctctgatgccaaactacaatcccggtctacatgtacgagaagagacgagatgacagac 1272

Db 1032 GTCCGACACGCCCTTACGACAGCAGCGCCAGCTACGAGAAGGAGAAGAAATGATGAAGTC 1091

Qy 1273 ccgatgatgaccaaaagccatcaataacgacctcagcttatctaggggctgaagccttcg 1332

Db 1092 CCACGTGATGACCAAGCCATCAACACGCCATCAACTACCTGGGGCCGAGTCCCTCGG 1151

Qy 1333 ccccttagtcagactcgcctcgtcccaacctctgagatggtccacgtcattcagcagctg 1392

Db 1152 CCGCTGTGTGACAGCGCCCGGG---CGGTTCGAGGTGTCCCGGTCTATCAGCCCGAT 1208

Qy 1393 gtaccatagcaacttactcggcgcatatgccaatggggggcccccgaggaga----- 1446

Db 1209 GTACCAAGCTGCACAAGCGCTCGCGGAGGACCCCGGCTCCAACCACTCGGCCAGGA 1268

Qy 1446 -----tggaataaagcggatctcctcctcgcagagaagatcttgccttctgaacagg 1497

Db 1269 CAGCGCGCTGAGAACCTGCTGCTCTCCAAGGCCAAGTTGGTGCCTCGGAGCGCA 1328

Qy 1498 tctgtcccccataaactggtccccgagctcacagacacgacagcaaccacgagg---a 1554

Db 1329 GGGTCCCGGAGCAACAGCTGTCAAGACTCCACGGACACCGAGACCAACACGAGGAGCA 1388

Qy 1555 tcgcaacatctctaccagcaaaacacgtgttctctcccccagcccgcaatggatgcc 1614

Db 1389 GCGCAGCGGTCTCATCTACCTGACCAACACCATCGCCCGCACGCGCAACGGCTTGT 1448

Qy 1615 tctctgaagaggtccctcgtcttttgaactcctcctcaagccccctccctcgtcctgag 1674

Db 1449 ---GCTCAAGAGGAGGACCGCGCTACGACCTGCTGCGCGCCGCTCCGAGAACTCGCA 1505

Qy 1675 ggactccatcaaatgatcaaaaaggagggtgagtgatgtgttctcgatgtgacca 1734

Db 1506 GGACGCGCTCGCGTGTGTCACACACGCGGGAGCAGATGAAGGTGTACAAAGTGGACACA 1565

Qy 1735 ctgcacgctctcttcttagattatgtgatgttccacatccacatggggtgacctggttt 1794

Db 1566 CTGCGGGGTGCTCTCTCGGTATCAGTCTATACACCATCCACATGGGCTGCCACGGCTT 1625

Qy 1795 ccgtgacctcttctgagtgataacatgtgtggtatcgaagacacgacgctatgagttctc 1854

Db 1626 CCGTATCCTTTTGTAGTCAACATGTGGGTACCCACAGCCAGGACCGGTACGAGTTCTC 1685

Qy 1855 cttccacatccgagagagacacag 1881

Db 1686 CTCGACATAACGCGGGGAGCACCG 1712

RESULT 3

S80876

LOCUS S80876 1788 bp mRNA PRI 27-MAR-1997

DEFINITION IKAROS=hik1 (alternatively spliced) [human, Jurkat T cell line, mRNA partial, 1788 nt].

ACCESSION S80876

NID g1911482

VERSION S80876.1 GI:1911482

KEYWORDS human Jurkat T cell line.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1788)

AUTHORS Molnar, A., Wu, P., Largespada, D.A., Vorkamp, A., Scherer, S., Copeland, N.G., Jenkins, N.A., Bruns, G. and Georgopoulos, K.

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 05:21:26 : Search time 701.47 Seconds
(without alignments)
8995.016 Million cell updates/sec

Title: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagcagcgacaccgctcgg.....gaactcaaacccctcgcg 1984

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_em.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sts.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_v1.*
- 18: em_fun.*
- 19: em_htg.*
- 20: em_hum1.*
- 21: em_hum2.*
- 22: em_in.*
- 23: em_om.*
- 24: em_or.*
- 25: em_ov.*
- 26: em_pat.*
- 27: em_ph.*
- 28: em_pl.*
- 29: em_ro.*
- 30: em_sts.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_v1.*
- 34: gb_htg1.*
- 35: gb_htg2.*
- 36: gb_in1.*
- 37: gb_in2.*
- 38: em_ba1.*
- 39: em_ba2.*
- 40: em_hum3.*
- 41: em_hum4.*
- 42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description
1	1521	76.7	1521	12	AF001293 Mus muscu

2	428.2	21.6	3629	11	HSU40462
3	411.8	20.8	1788	10	S80876
4	399.6	20.1	1557	4	GGIKTRF
5	348	17.5	1550	12	MUSIKAROS
6	335.6	16.9	2496	4	OMU92201
7	331.4	16.7	1847	12	AF044257
8	290.6	14.6	2183	4	OMU92200
9	260.6	13.1	2309	4	AF092175
10	226	11.4	2688	12	AB017615
11	208.6	10.5	2301	4	OMU92198
12	187.2	9.4	684	12	S74708
13	166.2	8.4	415	4	XLUS92202
14	164	8.3	2079	4	OMU92199
15	145.6	7.4	986	4	AF024439
16	111.6	5.6	148640	4	AF056116
17	91.8	4.6	3369	10	HSU411806
18	91.8	4.6	3186	10	HSOZF
19	90.8	4.6	3376	3	BT02FGENE
20	90	4.5	2686	36	SPU19831
21	89.4	4.5	2582	9	HUMZIF1
22	89.4	4.5	4208	11	AF011573
23	89.4	4.5	2582	14	G28565
24	89.2	4.5	3923	10	HSU57796
25	89.2	4.5	4848	11	HUMLD512P3
26	88.6	4.5	2342	12	MMZFPP35
27	88.6	4.5	2268	12	MUSZFPPB3
28	87.2	4.4	1944	9	HUMZNF8
29	85.8	4.3	1600	4	MSSNABP
30	85.6	4.3	4194	9	AB018303
31	85.6	4.3	2432	11	HSU09413
32	85.6	4.3	3326	12	RNU41164
33	85.6	4.3	110474	35	AC007494
34	85.2	4.3	4445	4	XLXFIN
35	85	4.3	945	36	OCZF2
36	84.8	4.3	789	36	OVZF1
37	84.8	4.3	1896	36	OVZF12
38	84.6	4.3	1661	12	MMZ224763
39	84.4	4.3	2615	12	MMRR2
40	84.4	4.3	819	12	MUSFPB
41	84	4.2	1434	12	MMNZFP
42	84	4.2	3492	12	MMZFPP37P1
43	84	4.2	185032	35	AC007607
44	83	4.2	2094	12	MMZFPP39
45	82.6	4.2	5471	9	AB007886

ALIGNMENTS

RESULT 1

AF001293

LOCUS

DEFINITION

ACCESSION

NID

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

AF001293 1521 bp mRNA ROD 04-JUN-1997
Mus musculus transcription factor aiolos mRNA, partial cds.

AF001293
92150043

AF001293.1 GI:2150043

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1521)

Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T.,

Gonzales, E., Wu, P., Neben, S. and Georgopoulos, K.

Aiolos, a lymphoid restricted transcription factor that interacts

with Ikaros to regulate lymphocyte differentiation

EMBO J. 17, 2004-2013 (1997)

2 (bases 1 to 1521)

Morgan, B.

Direct Submission

Submitted (24-APR-1997) CBRC, MGH East, Bldg. 149 13th Street,

Charlestown, MA 02129, USA

Location/Qualifiers